MATHEMATISCHE ZEITSCHRIFT

UNTER STÄNDIGER MITWIRKUNG VON

E. KAMKE
TÜBINGEN

R. NEVANLINNA
HELSINKI

E. SCHMIDT
BERLIN

F. K. SCHMIDT
MÜNSTER

H. WIELANDT
MAINZ

HERAUSGEGEBEN VON

K. KNOPP
TÜBINGEN

WISSENSCHAFTLICHER BEIRAT:

W. BLASCHKE L. FEJÉR G. HERGLOTZ
A. E. INGHAM H. KNESER O. PERRON
W. SÜSS H. WEYL

53. BAND 2. HEFT

(ABGESCHLOSSEN AM 10. AUGUST 1950)

BERLIN · GÖTTINGEN · HEIDELBERG
SPRINGER-VERLAG
1950

Math.
Z.

53. Band. **Inhalt:** 2. Heft.

Seite

Sansone, Prof. G., Firenze (Italien), Via Francesco Crispi: La formula di approssimazione asintotica dei polinomi di Tchebychef-Laguerre col procedimento di J. V. Uspensky 97

Nevanlinna, Prof. Dr. R., Helsinki (Finnland), Pietarink 5a: Bemerkungen zur Lösbarkeit der ersten Randwertaufgabe der Potentialtheorie auf allgemeine Flächen 106

Kneser, Prof. H., Tübingen, Brunsstr. 31: Eine direkte Ableitung des Zornschen Lemmas aus dem Auswahlaxiom 110

Nöbeling, Prof. Dr. G., Erlangen, Lammerstr. 5: Über die metrische Struktur eines Kontinuums endlicher Länge im Euklidischen Raum 114

Haupt, Prof. Dr. O., Erlangen, Spardorferstr. 45: Wronskische Determinante und lineare Abhängigkeit 122

Bompiani, Prof. Dr. E., Roma (Italien), Via Verona 22: Über die Gaußsche Krümmung der Überraumflächen 131

Pickert, Doz. Dr. G., Tübingen, Sonnenstr. 2: Eine Normalform für endliche rein-inseparable Körpererweiterungen 133

Tricomi, Prof. F., California Institute of Technology, Pasadena 4, Calif. (USA.): Asymptotische Eigenschaften der unvollständigen Gammafunktion 136

Collatz, Prof. Dr. H., Hannover, Tizianstr. 1: Über die Konvergenzkriterien bei Iterationsverfahren für lineare Gleichungssysteme 149

Lorenzen, Doz. Dr. P., Bonn a. Rh., Luisenstr. 3: Konstruktive Begründung der Mathematik 162

Unveränderter Nachdruck 1975

Springer-Verlag Berlin Heidelberg New York

Mathematische Zeitschrift, Band 53, Heft 2, S. 97—105 (1950).

La formula di approssimazione asintotica dei polinomi di Tchebychef-Laguerre col procedimento di J. V. Uspensky.

Herrn Erich Kamke zum 60. Geburtstag am 18. August 1950 gewidmet.

Nota di

Giovanni Sansone a Firenze.

1. E' noto che per l'n. esimo polinomio $L_n^{(\alpha)}(x)$ di Tchebychef-Laguerre

$$(1) \qquad L_n^{(\alpha)}(x) = \frac{x^{-\alpha} e^x}{n!} \frac{d^n}{dx^n} [x^{n+\alpha} e^{-x}] \qquad (n = 1, 2, \ldots),$$

fissato α, per x variabile in un intervallo finito (a, b), $0 < a \leqq x \leqq b$, L. Fejér [1]) stabilí la formula di rappresentazione asintotica

$$(2) \quad L_n^{(\alpha)}(x) = \frac{\Gamma(n+\alpha+1)}{\sqrt{\pi}\, n!} \frac{2^{\alpha+1/2} e^{x/2}}{[(4n+1)x]^{[\alpha+1/2]/2}} \left\{ \cos\left[\sqrt{(4n+1)x} - \frac{2\alpha+1}{4}\pi \right] + O\left(\frac{x^{3/2}}{\sqrt{4n+1}} \right) \right\}$$

dalla quale, valutando il rapporto $\Gamma(n+\alpha+1)/n!$ con la formula di Stirling si deduce che vale uniformemente in (a, b) la formula

$$(3) \qquad L_n^{(\alpha)}(x) = O\left[\frac{e^{x/2}}{\sqrt{x}} \left(\frac{n}{x} \right)^{\alpha/2 - 1/4} \right].$$

E. Kogbetliantz [2]) osservò che dai ragionamenti di O. Perron [3]) segue che la stessa formula (3) vale per $x < n^{1/3-\varepsilon}$, $0 < \varepsilon < 1/3$, e la estese al caso $x \leqq \varkappa n$. Successivamente G. Ottaviani [4]) precisò che se $\alpha \geqq -1$ sussiste la (3) per $1/n \leqq x \leqq \varkappa n$, e se $\alpha < -1$ per $0 < a \leqq x \leqq \varkappa n$ (a = costante), con $\varkappa < 4$, e che entro tali limiti non può esistere un'approssimazione migliore.

---

[1]) L. Fejér, Über die Bestimmung asymptotischer Werte, Mathematikai és természettudományi értesitö, 27 (1909), pp. 1—33.

[2]) E. Kogbetliantz, Sur la sommabilité des series d'Hermite, Ann. Sc. de l'Éc. Norm. Sup., (3), 49 (1932), (pp. 137—221), p. 149.

[3]) O. Perron, Über das Verhalten einer ausgearteten hypergeometrischen Reihe bei unbegrenztem Wachstum eines Parameters, Journ. für die r. und ang. Math. 151 (1921), pp. 63—78.

[4]) G. Ottaviani, Sulla convergenza e sommabilità delle serie di Hermite: Disuguaglianze fondamentali per i polinomi di Laguerre e di Hermite; Ann. Sc. Norm. Sup. di Pisa, (2), 7 (1938), (pp. 1—29), p. 13.

La (3) può dedursi dalla seguente formula (cfr. n. 4, b)

$$(4)\quad L_n^{(\alpha)}(x) = \frac{\Gamma(n+\alpha+1)}{\sqrt{\pi}\, n!}\, \frac{2^{\alpha+1/2}\, e^{x/2}}{[(4n+1)x]^{(\alpha+1/2)/2}} \left\{ \cos\left[\sqrt{(4n+1)\,x} - \frac{2\alpha+1}{4}\pi\right] + \frac{\frac{x^2}{6} - \left(\alpha+\frac{1}{2}\right)x - \frac{1}{2}\left(\alpha^2 - \frac{1}{4}\right)}{\sqrt{(4n+1)x}} + \frac{R(n, x, \alpha)}{4n+1} \right\},$$

dove $R(n, x, \alpha)$ indica una quantità uniformemente limitata rispetto ad n quando x varia in (a, b), $0 < a \leqq x \leqq b$, dimostrata da J. V. Uspensky [5]) partendo dalla relazione integrale

$$(5)\quad L_n^{(\alpha)}(x) = \frac{(-1)^n\, \Gamma(n+\alpha+1)}{\sqrt{\pi}\,(2n)!\,\Gamma(\alpha+1/2)} \int_0^{\pi} H_{2n}(\sqrt{x}\,\cos\varphi)\, \operatorname{sen}^{2\alpha}\varphi\, d\varphi \quad (\alpha > -1/2),$$

tra $L_n^{(\alpha)}(x)$ e i polinomi $H_n(x)$ di Tchebychef-Hermite definiti con la relazione

$$(6)\qquad H_n(x) = e^{x^2} \frac{d^n e^{-x^2}}{dx^n} \qquad (n = 1, 2, \ldots).$$

La dimostrazione di J. V. Uspensky procede direttamente, senza ricorso alle proprietà delle funzioni di Bessel strettamente collegate ai polinomi $L_n^{(\alpha)}(x)$ [6]).

In questa nota noi riprendiamo il procedimento di J. V. Uspensky perché riteniamo che esso debba essere completato (cfr. n. 3) per stabilire che nella (4) $R(n, x, \alpha)$ è uniformemente limitata rispetto ad n, e per procurarci una migliore precisazione del termine complementare di una formula di Uspensky (cfr. n. 2) la quale consente una rappresentazione asintotica dei polinomi $L_n^{(\alpha)}(x)$ per $1/(4n+1) \leqq x \leqq (4n+1)^{2/3-\varepsilon}$, $0 < \varepsilon < 1/3$, (cfr. n. 4, a).

2. *Nell'integrale*

$$(7)\qquad I = \int_{-1}^{1} e^{m^3 x^2/2} f(mx)\, (1-x^2)^{\alpha-1/2}\, e^{i\lambda m x}\, dx$$

sia $\alpha > -1/2$, $m > 0$, $\lambda \geqq 1$, $\lambda m \geqq 1$; $(1-x^2)^{\alpha-1/2}$ *sia il valore principale della potenza, ed* $f(x)$ *indichi un polinomio di grado* p *in* x. *Si*

[5]) J. V. Uspensky, On the development of arbitrary functions in series of Hermite's and Laguerre's polynomials; Ann. of Math., (2), 28 (1927), (pp. 593—619), p. 609.

[6]) Cfr. ad es. F. Tricomi: a) Sviluppo dei polinomi di Laguerre e di Hermite in serie di funzioni di Bessel, Giorn. Ist. It. Attuari, 12 (1941), pp. 14—33; b) Sul comportamento asintotico dell'nesimo polinomio di Laguerre nell'intorno dell'ascissa $4n$; Comm. Math. Helvetici, 22 (1949), pp. 150—167. (Aggiunta fatta durante la correzione delle bozze) Cfr. anche la recente interessante memoria di F. Tricomi: Sul comportamento asintotico dei polinomi di Laguerre, Ann. di Mat. pura ed appl., (4), 28 (1949), pp. 263—289.

ha allora

$$(8)\quad I = \int_0^{\pi} e^{(m^2/2)\cos^2\varphi} f(m\cos\varphi)\, e^{i\lambda m\cos\varphi}\, \operatorname{sen}^{2\alpha}\varphi\, d\varphi$$

$$= \frac{2^{\alpha+1/2}\,\Gamma(\alpha+1/2)\,e^{m^2/2}}{(\lambda m)^{\alpha+1/2}} \left\{ \frac{f(m)+f(-m)}{2} \cos\left[\lambda m - (2\alpha+1)\,\pi/4\right] \right.$$

$$+ i\,\frac{f(m)-f(-m)}{2} \operatorname{sen}\left[\lambda m - (2\alpha+1)\,\pi/4\right]$$

$$+ \frac{F(-m)-F(m)}{2\lambda} \operatorname{sen}\left[\lambda m - (2\alpha+1)\,\pi/4\right]$$

$$\left. + i\,\frac{F(m)+F(-m)}{2\lambda} \cos\left[\lambda m - (2\alpha+1)\,\pi/4\right] + \frac{r}{\lambda^2 m^2} \sum_{\varkappa=0}^{p+3} q_\varkappa(\alpha)\, m^\varkappa \right\},$$

con $|r| < 1$, *le* $q_\varkappa(\alpha)$ *costanti positive (non negative) dipendenti unicamente da* α, *ed* $F(m)$ *data dalla formula*

$$(9)\qquad F(m) = \frac{\{m f'(m) + [m^2 + (2\alpha - 1)/4]\, f(m)\}\,(\alpha + 1/2)}{m}\ ^{7}).$$

Si consideri nel piano (x, u) la semistriscia limitata dal segmento $(-1, 1)$ dell'asse reale e dalle due semirette del semipiano $u \geqq 0$ di ascissa $x = -1$, $x = 1$; applicando il teorema integrale di Cauchy si ottiene

$$(10)\qquad I = I_1 + I_2$$

con

$$(11_1)\quad I_1 = \frac{2^{\alpha-1/2}}{(\lambda m)^{\alpha+1/2}}\, e^{m^2/2 - i[\lambda m - (2\alpha+1)\pi/4]}$$

$$\times \int_0^{\infty} e^{-imu/\lambda - u^2/2\lambda^2} e^{-u} f\left[-m + iu/\lambda\right] u^{\alpha-1/2} (1 - iu/2\lambda m)^{\alpha-1/2}\, du,$$

$$(11_2)\quad I_2 = \frac{2^{\alpha-1/2}}{(\lambda m)^{\alpha+1/2}}\, e^{m^2/2 + i[\lambda m - (2\alpha+1)\pi/4]}$$

$$\times \int_0^{\infty} e^{imu/\lambda - u^2/2\lambda^2} e^{-u} f\left[m + iu/\lambda\right] u^{\alpha-1/2} (1 + iu/2\lambda m)^{\alpha-1/2}\, du.$$

Si ha con notazioni evidenti

$$e^{imu/\lambda} = \cos\frac{mu}{\lambda} + i \operatorname{sen}\frac{mu}{\lambda}$$

$$= 1 - \frac{m^2 u^2}{2\lambda^2} \cos\frac{m\bar{u}}{\lambda} + i\,\frac{mu}{\lambda} - i\,\frac{m^3 u^3}{6\lambda^3} \operatorname{sen}\frac{m\bar{u}}{\lambda},$$

---

[7]) Cfr. J. V. Uspensky, mem. cit. in [5]), p. 605. La forma del termine complementare $\frac{r}{\lambda^2 m^2} \sum_{\varkappa=0}^{p+3} q_\varkappa(\alpha)\, m^\varkappa$ è nuova.

$$(12_1)\quad e^{imu/\lambda} = 1 + i\frac{mu}{\lambda} + \frac{u^2}{2\lambda^2} g_1(m, u, \lambda), \qquad |g_1(m, u, \lambda)| \leqq m^2\left(1 + \frac{mu}{3\lambda}\right);$$

$$(12_2)\qquad e^{-u^2/2\lambda^2} = 1 - \frac{u^2}{2\lambda^2} g_2(u, \lambda), \qquad |g_2(u, \lambda)| < 1;$$

$$(12_3)\qquad f[m + iu/\lambda] = f(m) + \frac{iu}{\lambda} f'(m) + \frac{u^2}{2\lambda^2} g_3(u, \lambda, m),$$

$$|g_3(u, \lambda, m)| \leqq \sum_{x=2}^{p} \frac{2u^{x-2}}{x!\,\lambda^{x-2}} |f^{(x)}(m)|.$$

Si osservi ora che per $x > 0$ per il valore principale della potenza $(1 + ix)^\mu$ si ha

$$(13)\quad (1 + ix)^\mu = (1 + x^2)^{\mu/2} e^{i\mu \operatorname{arctg} x}, \qquad (x > 0,\ 0 < \operatorname{arctg} x < \pi/2),$$

ma $(1 + x^2)^{\mu/2} = 1 + (\mu/2) x^2 (1 + \theta_1 x^2)^{\mu/2 - 1}$, $0 < \theta_1 < 1$, e per $\mu < 2$ è $(1 + \theta_1 x^2)^{\mu/2 - 1} \leqq 1$, e per $\mu > 2$, $(1 + \theta_1 x^2)^{\mu/2 - 1} < (1 + x^2)^{\mu/2 - 1}$, perciò

$$(14)\quad (1 + x^2)^{\mu/2} = 1 + \theta_2 \frac{|\mu|}{2} (1 + x^2)^{\mu/2 - 1} \quad \text{con} \quad 0 < |\theta_2| < 1.$$

E' pure per la (12_1), $[0 < \operatorname{arctg} x < x]$

$$e^{i\mu \operatorname{arctg} x} = 1 + i\mu \operatorname{arctg} x + \frac{\mu^2}{2} \theta_3 \left(1 + \frac{|\mu|\, x}{3}\right), \qquad 0 < |\theta_3| < 1,$$

$$\operatorname{arctg} x = x - x^3 \frac{\overline{\theta}_3}{(1 + \overline{\theta}_3 x^2)^2} = x + \theta_4 x^3, \qquad 0 < \overline{\theta}_3,\ |\theta_4| < 1,$$

$$(15)\quad e^{i\mu \operatorname{arctg} x} = 1 + i\mu x + \theta_5 \frac{\mu^2 x^2}{2} \left\{ \frac{2x}{|\mu|} + 1 + \frac{|\mu|\, x}{3} \right\}, \qquad 0 < |\theta_5| < 1,$$

e perciò dalle (13), (14), (15)

$$(16)\quad (1 + ix)^\mu = \left[1 + i\mu x + \theta_5 \frac{\mu^2 x^2}{2} \left\{ \frac{2x}{|\mu|} + 1 + \frac{|\mu|\, x}{3} \right\}\right] \left[1 + \theta_2 \frac{|\mu|}{2} x^2 (1 + x^2)^{\mu/2 - 1}\right],$$

e in particolare, indicando con ν_1 un numero intero non inferiore a $|(\alpha - 1/2)/2 - 1|$

$$(1 + iu/2\lambda m)^{\alpha - 1/2} =$$

$$\left[1 + i\left(\alpha - \frac{1}{2}\right) \frac{u}{2\lambda m} + \theta_5 \left|\alpha - \frac{1}{2}\right| \frac{u^2}{3\lambda^2 m^2} \left\{ \frac{u}{\lambda m} + \left|\alpha - \frac{1}{2}\right| + \left(\alpha - \frac{1}{2}\right)^2 \frac{u}{6\lambda m} \right\}\right]$$

$$\times \left[1 + \frac{1}{2}\left|\alpha - \frac{1}{2}\right| \theta_2 \frac{u^2}{4\lambda^2 m^2} \left(1 + \frac{u^2}{4\lambda^2 m^2}\right)^{\nu_1}\right],$$

e perciò per $\lambda m \geqq 1$

$$(12_4)\qquad (1 + iu/2\lambda m)^{\alpha - 1/2} = 1 + i\left(\alpha - \frac{1}{2}\right) \frac{u}{2\lambda m} + \frac{u^2}{\lambda^2 m^2} g_4(u, \lambda, m)$$

$$|g_4(u, \lambda, m)| \leqq g(u)$$

dove $g(u)$ è un polinomio in u, a coefficienti non negativi, dipendenti unicamente da α, di grado $2\nu_1 + 3$.

Posto

$$\varphi(u) = e^{i m u/\lambda - u^2/2\lambda^2} f[m + iu/\lambda](1 + iu/2\lambda m)^{\alpha - 1/2} \tag{17}$$

si ha

$$\varphi(0) = f(m), \quad \varphi'(0) = i[m f'(m) + \{m^2 + (2\alpha - 1)/4\} f(m)]/\lambda m \tag{18}$$

e per le $(12_1), \ldots, (12_4)$ nelle ipotesi $\lambda m \geqq 1$, $\lambda \geqq 1$ si trova

$$\varphi(u) = \varphi(0) + u\varphi'(0) + \tau \frac{u^2 \omega(u)}{2\lambda^2}, \qquad |\tau| < 1, \tag{19}$$

$$\omega(u) = \frac{1}{m^2} \sum_{\varkappa=0}^{p+3} \varphi_\varkappa(u)\, m^\varkappa, \tag{20}$$

dove le $\varphi_\varkappa(u)$ indicano polinomi in u, a coefficienti non negativi, dipendenti unicamente da α, di grado non superiore a $2\nu_1 + p + 8$.

L'integrale (11_2) per le (17) e (19) diventa

$$I_2 = \frac{2^{\alpha - 1/2}}{(\lambda m)^{\alpha + 1/2}} e^{m^2/2 + i[\lambda m - (2\alpha + 1)\pi/4]} \times \int_0^\infty e^{-u} u^{\alpha - 1/2} \left[\varphi(0) + u\varphi'(0) + \tau \frac{u^2 \omega(u)}{2\lambda^2}\right] du\,;$$

ma

$$\int_0^\infty e^{-u} u^{\alpha - 1/2}\, du = \Gamma(\alpha + 1/2), \qquad \int_0^\infty e^{-u} u^{\alpha + 1/2}\, du = (\alpha + 1/2)\,\Gamma(\alpha + 1/2),$$

$$\frac{1}{2} \int_0^\infty e^{-u} u^{\alpha + 3/2} \omega(u)\, du = \frac{1}{2m^2} \sum_{\varkappa=0}^{p+3} m^\varkappa \int_0^\infty e^{-u} u^{\alpha + 3/2} \varphi_\varkappa(u)\, du$$

$$= \frac{\Gamma(\alpha + 1/2)}{2m^2} \sum_{\varkappa=0}^{p+3} \bar{q}_\varkappa(\alpha)\, m^\varkappa,$$

con le $\bar{q}_\varkappa(\alpha)$ *costanti non negative, unicamente dipendenti da* α, e ne risulta

$$I_2 = \frac{2^{\alpha - 1/2}\,\Gamma(\alpha + 1/2)}{(\lambda m)^{\alpha + 1/2}} e^{m^2/2 + i[\lambda m - (2\alpha + 1)\pi/4]} \times \left[f(m) + i \frac{[m f'(m) + \{m^2 + (2\alpha - 1)/4\} f(m)]\,(\alpha + 1/2)}{\lambda m} + \frac{r_1}{2\lambda^2 m^2} \sum_{\varkappa=0}^{p+3} \bar{q}_\varkappa(\alpha)\, m^\varkappa\right], \qquad |r_1| < 1,$$

ed operando analogamente sull'espressione di I_1 data dalla (11_1) si ricava appunto la (8).

3. Ci occorre un'altra limitazione dell'integrale (7) nel caso che $f(x)$ *sia una trascendente intera che soddisfa la limitazione* (u, v reali)

$$|f(u + iv)| \leqq \frac{e^{\lambda |v|}}{\lambda^{l+1}} P_{6l+11}[(u^2 + v^2)^{1/4}]\, e^{(u^2 + v^2)^{1/2}/\lambda} \tag{21}$$

dove $P_{6l+11}[(u^2+v^2)^{1/4}]$ *è un polinomio della variabile* $(u^2+v^2)^{1/4}$ *a coefficienti positivi o nulli, indipendenti da* λ *e da* m, *di grado* $6l+11$, *e l'intero* l *soddisfa la limitazione*

$$l \geqq \frac{1}{3\tau}\left[\frac{13}{2}\left(\frac{1}{3}-\tau\right)+\nu\left(\frac{4}{3}-\tau\right)\right] \tag{22}$$

essendo

$$\alpha > -1/2, \quad m > 0, \quad \lambda \geqq 1, \quad 1/\lambda \leqq m \leqq \lambda^{1/3-\tau}, \quad 0 < \tau < \varepsilon < 1/3, \tag{23}$$

e ν *un numero non inferiore a* $2|\alpha - 1/2|$.

Vogliamo dimostrare che *in queste ipotesi esiste una costante* C *indipendente da* λ *e da* m *tale che*

$$|I| \leqq C\,\frac{2^{\alpha+1/2}\,\Gamma(\alpha+1/2)\,e^{m^2/2}}{(\lambda m)^{\alpha+1/2}}. \tag{24}$$

Consideriamo la poligonale $C(T)$, $T = \lambda^{1/3-\tau}/m$, con i vertici consecutivi nei punti $-1, -1+iT, 1+iT, 1$. La poligonale $C(T)$ insieme al segmento di estremi 1 e -1 determina un circuito chiuso e applicando il teorema integrale di Cauchy si ottiene che I si scinde nella somma dei tre termini

$$(25_1)\quad e^{-\lambda m T}\int_{-1}^{1} e^{m^2(x^2-T^2)/2}\,e^{i\,m^2 x T}\,f[m\,x+i\,m\,T]\,[1-(x+iT)^2]^{\alpha-1/2}\,e^{i\,\lambda m x}\,dx,$$

$$(25_2)\quad \frac{2^{\alpha-1/2}}{(\lambda m)^{\alpha+1/2}}\,e^{m^2/2-i[\lambda m-(2\alpha+1)\pi/4]} \times \int_0^{\lambda^{4/3-\tau}} e^{-i\,m\,u/\lambda-u^2/2\lambda^2}\,e^{-u}\,f[-m+i\,u/\lambda]\,u^{\alpha-1/2}(1-i\,u/2\,\lambda\,m)^{\alpha-1/2}\,du,$$

$$(25_3)\quad \frac{2^{\alpha-1/2}}{(\lambda m)^{\alpha+1/2}}\,e^{m^2/2+i[\lambda m-(2\alpha+1)\pi/4]} \times \int_0^{\lambda^{4/3-\tau}} e^{i\,m\,u/\lambda-u^2/2\lambda^2}\,e^{-u}\,f[m+i\,u/\lambda]\,u^{\alpha-1/2}(1-i\,u/2\,\lambda\,m)^{\alpha-1/2}\,du.$$

Essendo per $|x| \leqq 1$

$$m^2(x^2-T^2) \leqq m^2 - m^2T^2 \leqq \lambda^{2/3-2\varepsilon} - \lambda^{2/3-2\tau} = -\lambda^{2/3-2\tau}(1-\lambda^{-2(\varepsilon-\tau)}),$$

$$(m^2x^2+m^2T^2)^{1/4} \leqq (m^2+m^2T^2)^{1/4} \leqq (\lambda^{2/3}+\lambda^{2/3})^{1/4} = 2^{1/4}\lambda^{1/6},$$

$$(m^2x^2+m^2T^2)\,mT/\lambda \leqq 2\,\lambda^{2/3+1/3-\tau-1} = 2/\lambda^{\tau},$$

$$T \leqq \lambda^{1/3-\tau}/(1/\lambda) < \lambda^{4/3}, \quad T \geqq \lambda^{1/3-\tau}/\lambda^{1/3-\varepsilon} > 1,$$

$$|1-(x+iT)^2|^{\alpha-1/2} \leqq [1+(1+T)^2]^{|\alpha-1/2|} \leqq [1+(1+\lambda^{4/3})^2]^{|\alpha-1/2|},$$

ne viene, tenuto conto della (21), che il termine (25_1) è maggiorato da

$$\frac{2}{\lambda^{l+1}}\,e^{-\lambda^{2/3-2\tau}(1-\lambda^{-2(\varepsilon-\tau)})}\,P_{6l+11}[2^{1/4}\lambda^{1/6}]\,e^{2/\lambda^{\tau}}\,[1+(1+\lambda^{4/3})^2]^{|\alpha-1/2|}$$

che per $\lambda \to \infty$ è un infinitesimo esponenziale, e tale si mantiene moltiplicandolo per $(\lambda m)^{\alpha+1/2}/e^{m^2/2} < \lambda^{(4/3-\varepsilon)|\alpha+1/2|}$.

Poiché per $u < \lambda^{4/3-\tau}$ si ha

$$m^2 + \frac{u^2}{\lambda^2} \leqq \lambda^{2/3-2\varepsilon} + \lambda^{2/3-2\tau} < 2\lambda^{2/3-2\tau}$$

$$P_{6l+11}\left[\left(m^2 + \frac{u^2}{\lambda^2}\right)^{1/4}\right] \leqq C_1 \lambda^{(1/3-\tau)(6l+11)/2}$$

$$\left(m^2 + \frac{u^2}{\lambda^2}\right)\frac{u}{\lambda} \leqq 2\lambda^{2/3-2\tau}\lambda^{-2/3-\tau} = 2\lambda^{-3\tau} \leqq 2,$$

il termine (25_2), a motivo della (22) non supera il prodotto di $2^{\alpha+1/2}e^{m^2/2}(\lambda m)^{-(\alpha+1/2)}$ per la somma

$$C_1 \frac{e^2}{2}\lambda^{(1/3-\tau)(6l+11)/2}\int_0^1 u^{\alpha-1/2}(1+u^2/4)^{|\alpha-1/2|/2}\,du$$

$$+ C_2 \lambda^{(1/3-\tau)(6l+11)/2+(4/3-\tau)(\nu+1)}/\lambda^{l+1}$$

con C_2 costante come la C_1.

Per la (22) si ha $(1/3-\tau)(6l+11)/2 + (4/3-\tau)(\nu+1) \leqq l+1$, e perciò per il termine (25_2) vale una limitazione come la (24); un'analoga limitazione vale per il termine (25_3), e la (24) risulta così dimostrata.

4. a) Siamo ora in grado di pervenire alla desiderata formula di rappresentazione asintotica dei polinomi di Tchebychef-Laguerre.

Abbiamo dimostrato [8]) che per il polinomio $H_{2n}(x)$ di Tchebychef-Hermite per ogni intero non negativo l sussiste la formula di approssimazione asintotica

$$(26)\qquad H_{2n}(x) = \pi^{1/4}2^n[(2n)!]^{1/2}e^{x^2/2}$$

$$\times\left\{f_{2n}(0)\cos[N^{1/2}x] + \frac{f_{2n}(0)}{N^{1/2}}\frac{x^3}{6}\operatorname{sen}[N^{1/2}x] + \frac{t(2n,x)}{N}\right\},$$

$$f_{2n}(0) = (-1)^n\pi^{-1/4}[(2n-1)!!/(2n)!!]^{1/2}, \qquad N = 4n+1,$$

dove $t(2n, x)$ è una trascendente intera in x tale che

$$(27)$$

$$t(2n,x) = \frac{1}{N^{1/4}}\{P^{(1)}_{3l+3}(x)\cos[N^{1/2}x] + P^{(2)}_{3l+3}(x)\operatorname{sen}[N^{1/2}x]\} + \frac{\bar{t}(2n,x)}{N^{(l+1)/2}},$$

con $P^{(1)}_{3l+3}(x)$, $P^{(2)}_{3l+3}(x)$ polinomi in x di grado $3l+3$ con coefficienti limitati rispetto ad n,

$$(28)\qquad \bar{t}(2n,x) = \bar{t}_1(2n,x)\cos[N^{1/2}x] + \bar{t}_2(2n,x)\operatorname{sen}[N^{1/2}x]$$

dove $\bar{t}_1(2n,x)$, $\bar{t}_2(2n,x)$ sono due trascendenti intere in x che soddis-

---

[8]) G. Sansone, Due semplici limitazioni nel campo complesso delle funzioni associate ai polinomi di Tchebychef-Hermite e del termine complementare della loro rappresentazione asintotica; Math. Zeitschr., 52 (1950), pp. 593—598, formule (15), (16), (17).

fano entrambe la limitazione (u, v reali)

$$(29)\quad |\bar{t}_i(2n, u+iv)| \leqq \frac{e^{N^{1/2}|v|}}{N^{1/4}}\left[\sum_{r=0}^{6l+11} c_r (u^2+v^2)^{r/4}\right] e^{N^{-1/2}(u^2+v^2)|v|} \quad (i=1,2),$$

essendo le costanti c_r non negative, dipendenti unicamente da l e indipendenti da x.

Dalle (5) e (26) si ottiene

$$(30)\quad L_n^{(\alpha)}(x) = \frac{\Gamma(n+\alpha+1)}{\sqrt{\pi}\, n!\, \Gamma(\alpha+1/2)} \int_0^{\pi} e^{(x/2)\cos^2\varphi}\left\{\cos\left[\sqrt{Nx}\,\cos\varphi\right] + \frac{x^{3/2}\cos^3\varphi}{6N^{1/2}}\operatorname{sen}\left[\sqrt{Nx}\,\cos\varphi\right] + \frac{1}{f_{2n}(0)}\,\frac{t(2n,\sqrt{x}\,\cos\varphi)}{N}\right\}\operatorname{sen}^{2\alpha}\varphi\, d\varphi\,;$$

dalla (8) posto una volta $m=\sqrt{x}$, $\lambda = N^{1/2}$, $f(x)\equiv 1$, $(p=0)$, supposto $(\lambda m \geqq 1)$, $x \geqq 1/N$, e uguagliando le parti reali dei due membri, e una seconda volta $m=\sqrt{x}$, $\lambda = N^{1/2}$, $(x \geqq 1/N)$, $f(m)\equiv m^3$, $(p=3)$, uguagliando i coefficienti dell'immaginario si ricava che

$$(31_1)\quad \int_0^{\pi} e^{(x/2)\cos^2\varphi}\left\{\cos\left[\sqrt{Nx}\cos\varphi\right] + \frac{x^{3/2}\cos^3\varphi}{6N^{1/2}}\operatorname{sen}\left[\sqrt{Nx}\cos\varphi\right]\right\}\operatorname{sen}^{2\alpha}\varphi\, d\varphi$$

$$= \frac{2^{\alpha+1/2}\,\Gamma(\alpha+1/2)}{[(4n+1)x]^{(\alpha+1/2)/2}}\left[\cos\left[\sqrt{(4n+1)x} - \frac{2\alpha+1}{4}\pi\right] + \frac{\frac{x^2}{6} - \left(\alpha+\frac{1}{2}\right)x - \frac{1}{2}\left(\alpha^2-\frac{1}{4}\right)}{\sqrt{(4n+1)x}}\operatorname{sen}\left[\sqrt{(4n+1)x} - \frac{2\alpha+1}{4}\pi\right] + \frac{r}{Nx}\sum_{\varkappa=0}^{6} q_{1,\varkappa}(\alpha)\, x^{\varkappa/2}\right], \qquad |r|<1\,;$$

si ha infine dalle (27), (28), (29), (8), (24), e osservando che

$$|1/f_{2n}(0)| = \sqrt{\pi}\,\sqrt[4]{4n+1}\,\sqrt[4]{(2n+1)/(4n+1)} < \sqrt{\pi}\, n^{1/4},$$

$$(31_2)\quad \left|\int_0^{\pi} e^{(x/2)\cos^2\varphi}\, t(2n, \sqrt{x}\,\cos\varphi)\operatorname{sen}^{2\alpha}\varphi\, d\varphi\right|$$

$$= \left|\int_0^{\pi} e^{(x/2)\cos^2\varphi}\left[\frac{1}{f_{2n}(0)\,N^{1/4}}\left\{\cos\left[\sqrt{Nx}\,\cos\varphi\right] P_{3l+3}^{(1)}(x^{1/2}\cos\varphi) + \operatorname{sen}\left[\sqrt{Nx}\,\cos\varphi\right] P_{3l+3}^{(2)}(x^{1/2}\cos\varphi)\right\} + \frac{1}{f_{2n}(0)\,N^{1/4}}\,\frac{N^{1/4}\,\bar{t}\,(2n, x^{1/2}\cos\varphi)}{N^{(l+1)/2}}\right]\operatorname{sen}^{2\alpha}\varphi\, d\varphi\right|$$

$$\leqq C\,\frac{2^{\alpha+1/2}\,\Gamma(\alpha+1/2)}{[(4n+1)x]^{(\alpha+1/2)/2}}\sum_{\varkappa=0}^{3(l+2)} q_{2,\varkappa}(\alpha)\, x^{\varkappa/2}, \qquad (C = \text{costante}),$$

e dalle (30), (31_1), (31_2) si ottiene che *per*

$$(32) \qquad \alpha > -1/2, \quad 1/(4n+1) \leqq x \leqq (4n+1)^{2/3-2\varepsilon}, \quad 0 < \varepsilon < 1/3$$

se ν *è un numero non inferiore a* $2|\alpha - 1/2|$, *e* $0 < \tau < \varepsilon$, *ad ogni* l *intero tale che*

$$(22) \qquad l \geqq \frac{1}{3\tau}\left[\frac{13}{2}\left(\frac{1}{3}-\tau\right)+\nu\left(\frac{4}{3}-\tau\right)\right]$$

corrisponde una formula di approssimazione asintotica dei polinomi $L_n^{(\alpha)}(x)$

$$(33) \quad L_n^{(\alpha)}(x) = \frac{\Gamma(n+\alpha+1)}{\sqrt{\pi}\, n!} \frac{2^{\alpha+1/2}\, e^{x/2}}{[(4n+1)x]^{(\alpha+1/2)/2}} \left\{ \cos\left[\sqrt{(4n+1)x} - \frac{2\alpha+1}{4}\pi\right] \right.$$
$$+ \frac{\frac{x^2}{6} - \left(\alpha+\frac{1}{2}\right)x + \frac{1}{2}\left(\alpha^2-\frac{1}{4}\right)}{\sqrt{(4n+1)x}} \operatorname{sen}\left[\sqrt{(4n+1)x} - \frac{2\alpha+1}{4}\pi\right]$$
$$\left. + \frac{R(n,x,\alpha)}{(4n+1)x} \sum_{\varkappa=0}^{3l+3} q_\varkappa(\alpha)\, x^{\varkappa/2} \right\}$$

dove le $q_\varkappa(\alpha)$ *sono costanti non negative, dipendenti unicamente da* α, *e fissato* α, $|R(n,x,\alpha)| \leqq 1$ *per*

$$1/(4n+1) \leqq x \leqq (4n+1)^{2/3-2\varepsilon} \qquad (n = 1, 2, \ldots).$$

b) Dalla (33) quando x varia in un intervallo finito (a, b), $0 < a \leqq x \leqq b$ si deduce la ricercata formula (4) di Uspensky.

5. Noi abbiamo dimostrato la (33) nel caso di $\alpha > -1/2$, ma essa si estende facilmente al caso di α reale qualunque.

Se ad es. $-3/2 < \alpha \leqq -1/2$, poiché [9])

$$\frac{d}{dx} L_n^{(\alpha)}(x) = -L_{n-1}^{(\alpha+1)}(x), \qquad \frac{d^2}{dx^2} L_n^{(\alpha)}(x) = L_{n-2}^{(\alpha+2)}(x)$$

e perciò

$$L_n^{(\alpha)}(x) = -\frac{x}{n} L_{n-2}^{(\alpha+2)}(x) + \frac{\alpha - x + 1}{n} L_{n-1}^{(\alpha+1)}(x),$$

calcolando il secondo membro con la (33) si ottiene una formula analoga salvo che ν indica un numero positivo tale che $\nu \geqq 2[(\alpha+2) - 1/2]$.

In generale *se* $\alpha < -1/2$ *e* $-(2s+1)/2 < \alpha < -(2s-1)/2$ $(s = 1, 2, \ldots)$, *ed* $n + \alpha + 1 > 0$ *sussiste la* (33), *salvo a fare* $\nu \geqq 2[(\alpha + 2s) - 1/2]$.

[9]) Cfr. ad es G. Sansone, Sviluppi in serie di funzioni ortogonali, (2ª ed., Bologna, 1946), p. 277, (11), e p. 278, (14_1).

(Eingegangen am 7. September 1949.)

Mathematische Zeitschrift, Band 53, Heft 2, S. 106—109 (1950).

Bemerkungen zur Lösbarkeit der ersten Randwertaufgabe der Potentialtheorie auf allgemeinen Flächen.

Herrn Erich Kamke zum 60. Geburtstag am 18. August 1950 gewidmet.

Von

Rolf Nevanlinna in Helsinki.

1. Die Uniformisierungstheorie öffnet eine prinzipielle Möglichkeit zur Lösung des potentialtheoretischen Grundproblems, eine harmonische Funktion mit vorgegebenen Randwerten auf einer gegebenen Fläche F zu konstruieren. Man bilde hierzu zu der Riemannschen Fläche F die universelle Überlagerungsfläche $\tilde{\Gamma}$. Diese gehört, mit Ausnahme gewisser trivialer Fälle, die uns hier nicht interessieren, zu dem hyperbolischen Typus, und man kann sie also auf das Innere des Einheitskreises $|z| < 1$ eineindeutig und konform abbilden. Jedem Punkt P der Fläche F wird dann eine abzählbare Punktmenge (z) entsprechen, welche mit Hilfe einer Gruppe von linearen Transformationen (S) zusammenhängen, die den Einheitskreis $|z| < 1$ in sich selbst überführen. Die Identifizierung der Punkte (z), welche in diesem Sinn äquivalent sind, führt zu einer Darstellung der gegebenen Fläche F.

2. Auch die Punkte ζ der Peripherie $|\zeta| = 1$ ordnen sich nach Äquivalenzgruppen (ζ), welche sämtliche, inbezug auf die Substitutionsgruppe (S) äquivalenten Punkte enthalten. Es sei l ein Radius des Kreises $|z| < 1$ und ζ sein Endpunkt; dem Weg l und seinen äquivalenten Wegen $S(l)$ entspricht auf der Fläche F ein Weg L. Kann man behaupten, daß L gegen den idealen Rand Γ von F strebt, in dem Sinn, daß die analytische Kurve L_r, welche dem in den Kreisring $r \leqq |z| < 1$ $(r < 1)$ fallenden Teil l_r des Radius l entspricht, für $r \to 1$ außerhalb jedes vorgegebenen kompakten Teils der Fläche F liegt? Die Antwort wird verschieden ausfallen, je nach den metrischen Verhältnissen des Randes Γ. Falls Γ das harmonische Maß Null hat, so wird der Weg i. A. nicht diese Konvergenzeigenschaft haben, sondern einen ergodischen Charakter besitzen. Anders verhält es sich bei den Flächen mit einem Rand von positivem Rand Γ. Dies ergibt sich aus folgender Überlegung.

3. Die Fläche F ist dann und nur dann „positiv berandet", wenn sie eine Greensche Funktion G hat. Verlegen wir den Pol dieser Funktion in den Punkt P_0, entsprechend der Punktgruppe $(z_1, z_2, \ldots)$ im Kreis

$|z|<1$, so wird

$$G=\sum_{\nu=1}^{\infty} g(z, z_\nu),$$

wo g die Greensche Funktion des Einheitskreises ist, eine Summe, die dann und nur dann konvergiert, wenn die Reihe

$$\sum_{\nu}(1-|z_\nu|)$$

konvergent ist. Diese letzte Bedingung ist somit notwendig und hinreichend dafür, daß die Greensche Funktion G existiert und der Rand Γ positives Maß hat[1]).

Nimmt man nun an, daß diese Konvergenzbedingung erfüllt ist, so hat G nach dem Fatouschen Satz radiale Randwerte fast überall auf der Peripherie $|z|=1$, und zwar sind diese Randwerte konstant gleich Null. Ist nun l ein entsprechender Radius, so muß der Weg L auf F die in Nr. 2 erwähnte Konvergenzeigenschaft besitzen; denn es ist $G \to 0$ auf L (für $|z| \to 1$), und L muß gegen den Rand Γ streben, da G in jedem kompakten Teil von Γ eine positive untere Grenze hat. Ich bezeichne jeden solchen Weg L einen Randweg der Fläche.

4. Wir haben gesehen, daß die obige Definition der Randwege L als Voraussetzung hat, daß der Rand Γ von positivem Maß ist, und hiermit ist eine Grundlage geschaffen für die Lösung der ersten Randwertaufgabe der Potentialtheorie für eine solche Fläche F.

Wollen wir zunächst annehmen, es existiere auf F eine eindeutige, beschränkte Potentialfunktion u. Sofern u nicht konstant ist, impliziert die Existenz einer solchen Funktion, daß der Rand positives Maß hat. Geht man nun durch die obige Fundamentalabbildung $P \to z$ zu dem veränderlichen z über, so wird u eine Funktion von z, die für $|z|<1$ regulär harmonisch und beschränkt ist; sie bleibt ferner für die Substitutionen S invariant. Nach dem Fatouschen Satz hat u fast überall auf der Peripherie $|z|=1$ wohlbestimmte radiale Grenzwerte; wegen des automorphen Charakters von u sind diese Grenzwerte ferner inbezug auf die Substitutionen S invariant, so daß jeder Punktgruppe $S(\zeta)$ ein und derselbe Grenzwert zugeordnet ist, mit Ausnahme einer Nullmenge auf der Peripherie $|\zeta|=1$, wo eventuell keine Randwerte vorhanden sind. Die Funktion u läßt sich mit Hilfe des Poissonschen Integrals durch diese Randwerte, welche eine im Lebesgueschen Sinn meßbare Funktion erklären, darstellen:

$$u(re^{i\varphi})=\frac{1}{2\pi}\int_0^{2\pi} u(e^{i\vartheta})\,\frac{1-r^2}{1+r^2-2r\cos(\vartheta-\varphi)}\,d\vartheta.$$

[1]) Vgl. meine Monographie: Eindeutige analytische Funktionen (Spenger, Berlin 1936).

5. Umgekehrt kann man nun jedem Randweg L einen beliebigen beschränkten Wert $f(L)$ zuordnen und eine Funktion u aufsuchen, die auf F harmonisch und beschränkt ist und auf „fast allen" Wegen L gegen den vorgegebenen Randwert f strebt. Ob die Lösung dieser Aufgabe gelingt, hängt davon ab, inwiefern die Funktion $f(e^{i\vartheta})$, welche durch Übertragung aus $f(L)$ entsteht, im Lebesgueschen Sinn meßbar ist. Vorausgesetzt, daß dem so ist, hat man für die Lösung des Randwertproblems die obige Darstellung mit Hilfe des Poissonschen Integrals. Sie stellt nämlich eine inbezug auf die Gruppe (S) automorphe Potentialfunktion mit den vorgegebenen Randwerten dar und ergibt also, auf die Fläche F übertragen, das gesuchte eindeutige Potential.

6. Nun ist aber die Meßbarkeitsfrage, zu der die obige Betrachtung führt, nicht leicht zu entscheiden. Es zeigt sich nämlich, daß schon das viel speziellere Problem: Die Randwegmenge (L) in zwei komplementäre Mengen $(L)_1$, $(L)_2$ zu zerlegen, welche beide positives Maß haben, Schwierigkeiten begegnet. Dieses besondere Problem hat für gewisse aktuelle Fragen der Theorie der offenen Riemannschen Flächen Interesse. Wie oben schon bemerkt wurde, existiert auf einer nullberandeten Fläche F keine nichtkonstante, eindeutige Potentialfunktion. Die Frage ist nun: gibt es auf jeder positivberandeten Fläche stets solche Funktionen? Man sieht den Zusammenhang dieser Frage mit der obigen Randwertaufgabe ein. In der Tat, eine notwendige und hinreichende Bedingung dafür, daß solche Potentialfunktionen existieren, ist gerade, daß die Menge (L) eine Zerlegung in zwei Teilmengen von positivem Maß zuläßt. Die Notwendigkeit dieser Bedingung ist eine direkte Folgerung aus den Überlegungen von Nr. 4. Die Bedingung ist aber auch hinreichend. Denn wenn eine solche Zerlegung möglich ist, so kann man den Wegen $(L)_1$ den Randwert $f = 1$, den Wegen $(L)_2$ die Werte Null zuordnen, und der obige Ansatz mit dem Poissonschen Integral führt zu einer beschränkten, nichtkonstanten, eindeutigen Potentialfunktion auf F.

Falls die Fläche F endliches Geschlecht hat, so läßt sich diese Konstruktion stets vornehmen, und dasselbe gilt, wie ich in einer demnächst erscheinenden Arbeit gezeigt habe, auch für gewisse Klassen von Flächen mit unendlichem Geschlecht, wobei ich allerdings den Beweis durch andere Methoden als die hier angedeutete gegeben habe.

7. Ich möchte zum Schluß durch ein Beispiel zeigen, daß die Frage nach der Meßbarkeit der hier betrachteten Punktmengen (ξ) auf der Peripherie $|z| = 1$ mit Vorsicht zu behandeln ist. Betrachten wir hierzu eine Fläche F, die entweder geschlossen ist oder einen Nullrand Γ hat. Wie in Nr. 2 bilde man mittels der Gruppe (S), welche der universellen Überlagerungsfläche $\tilde{F}$ von F entspricht, die Mengen (ξ) der äquivalenten Punkte auf der Peripherie. Dann ist eine der in Nr. 7 besprochenen Zerlegung entsprechende Einteilung der Peripheriepunkte

$\zeta = e^{i\vartheta}$ in zwei Teilmengen M_1 und M_2 vom positiven LEBESGUEschen Maß unmöglich. Denn würde eine solche Einteilung existieren, so daß die Punkte von M_1 nichtäquivalent sind mit den Punkten von M_2, so könnte man wieder das POISSONsche Integral mit den Randwerten 1 auf M_1, Null auf M_2 ansetzen, und dieses Integral würde dann ein nichtkonstantes Potential definieren, das auf der Fläche F eindeutig, regulär und beschränkt ist, was nach dem oben Gesagten (Nr. 4) unmöglich ist.

(Eingegangen am 3. Oktober 1949.)

Mathematische Zeitschrift, Band 53, Heft 2, S. 110—113 (1950).

Eine direkte Ableitung des Zornschen Lemmas aus dem Auswahlaxiom.

Herrn Erich Kamke zum 60. Geburtstag am 18. August 1950 gewidmet.

Von

Hellmuth Kneser in Tübingen.

Der Satz, den M. Zorn[1]) als ein allgemeines Prinzip der Mengenlehre aussprach, hat über die von Zorn schon damals angegebenen Anwendungen in der transfiniten Algebra hinaus wachsende Bedeutung für die verschiedensten Gebiete gewonnen. Wie schon Zorn bemerkte, folgt er aus dem Auswahlaxiom; eine Ableitung aus dem Auswahlaxiom ist aber meines Wissens erst an einer Stelle[2]) veröffentlicht worden, und zwar auf dem Wege über den Wohlordnungssatz. Da aber der Satz bei seinen Anwendungen grade den Wohlordnungssatz ersetzt, lohnt es sich, ihn ohne jenen Umweg zu beweisen. Der Beweis, den ich im folgenden gebe, schließt sich an Zermelos ersten Beweis des Wohlordnungssatzes an. Er benutzt den Begriff der Wohlordnung und führt zu einer etwas erweiterten Fassung des Zornschen Lemmas. Wendet man ihn zum Beweise des Wohlordnungssatzes an, so erhält man Zermelos ersten Beweis mit einer kleinen abkürzenden Wendung.

Natürlich kann man den Zornschen Satz aus dem Auswahlaxiom auch im Anschluß an Zermelos zweiten Beweis des Wohlordnungssatzes ableiten. Für die Darstellung der Mengenlehre in Bourbakis „Éléments", von der ja bisher nur der Auszug ohne Beweise vorliegt[3]), ist ein solcher Beweis vorgesehen, wie ich aus dem Kreise der Mitarbeiter des Werkes erfahre. Das Verhältnis der Beweise — wenn man vergleichen will — ist ähnlich wie bei den Zermeloschen: der „zweite" ist kunstvoller und eleganter; der „erste" geht grader aufs Ziel los, was auch der Grund sein mag, aus dem manche Lehrbücher ihm den Vorzug geben[4]).

In der Fassung von Bourbaki bezieht sich das Zornsche Lemma auf teilweise geordnete Mengen. Die ordnende Beziehung bezeichne ich mit $\leqq$, wenn die Gleichheit zugelassen, mit $<$, wenn sie ausgeschlossen sein soll. Das Lemma lautet:

[1]) Bull. Amer. math. Soc. 41 (1935), 667—670.

[2]) G. Birkhoff, Theory of Lattices, 2. Aufl. (1948).

[3]) N. Bourbaki, Éléments de mathématique I, 1, Théorie des ensembles (1939).

[4]) Z. B. A. Fraenkel, Einleitung in die Mengenlehre, 3. Aufl. (1928) und E. Kamke, Mengenlehre, 2. Aufl. (1947).

A. *Hat in einer nicht leeren teilweise geordneten Menge E jede nicht leere vollständig geordnete Teilmenge M eine obere Grenze*[5]*, so gibt es in E ein maximales Element*[5].

Wir setzen etwas weniger voraus und behaupten

B. *Hat in einer teilweise geordneten Menge E jede wohlgeordnete Teilmenge M eine obere Schranke $p(M)$, so gibt es in E ein maximales Element.*

Daß E nicht leer sei, brauchen wir nicht besonders vorauszusetzen, da ja jedenfalls $p(0) \in E$ ist. Die Abschwächung, daß statt der oberen Grenze nur eine obere Schranke vorausgesetzt ist, kann für die Anwendung des Satzes nützlich sein. Die andere Abschwächung, daß die obere Schranke nur bei wohlgeordneten Teilmengen, nicht bei allen vollständig geordneten vorausgesetzt wird, ist inhaltsmäßig unwichtig, da man aus jeder vollständig geordneten Menge eine wohlgeordnete Teilmenge herausgreifen kann, die mit ihr konfinal ist, d. h. genau dieselben oberen Schranken hat. Allerdings braucht man dazu noch einmal das Auswahlaxiom.

Satz B folgt durch eine einfache Anwendung des Auswahlaxioms aus

C. *In der teilweise geordneten Menge E habe jede wohlgeordnete Teilmenge M eine obere Schranke $p(M)$; ferner sei eine Abbildung $x \to fx$ von E in sich gegeben mit der Eigenschaft $x \leqq fx$. Dann gibt es in E ein Element w mit $w = fw$.*

Dieser Satz verhält sich zu dem „lemme fondamental“ bei Bourbaki[3]) wie Satz B zu Satz A.

Ist $x \in E$ nicht maximal, so ist die Menge der $y \in E$ mit $x < y$ nicht leer. Nach dem Auswahlaxiom können wir also jedem nicht maximalen x ein $fx > x$ zuordnen; für maximales x setzen wir $fx = x$. Sind nun in E die Voraussetzungen von B. erfüllt, so gelten mit fx die von C. Also gibt es nach C. ein $w = fw$, und dieses ist maximal. Wir brauchen also nur noch C. zu beweisen, und das geht ohne das Auswahlaxiom.

Anfangsstück einer teilweise geordneten Menge M heißt bekanntlich eine Menge $L \subseteq M$, wenn aus $x \in M$ und $x \leqq y \in L$ immer $x \in L$ folgt. Die Vereinigung einer Menge von Anfangsstücken von M ist wieder Anfangsstück von M.

Ist A ein echtes Anfangsstück der wohlgeordneten Menge M und y das erste nicht in A enthaltene Element von M, so besteht A aus allen $x \in M$ mit $x < y$, d. h. A ist der Abschnitt M_y von y in M.

Unter den Voraussetzungen von C. heiße eine Menge $L \subseteq E$ eine Kette, wenn L durch die teilweise Ordnung von E wohlgeordnet

[5]) Ein Element $z \in E$ heißt obere Schranke von $M \subseteq E$, wenn $x \leqq z$ für jedes $x \in M$ gilt. Obere Grenze ist die — wenn sie überhaupt vorhanden ist, eindeutig bestimmte — kleinste obere Schranke, d. h. eine obere Schranke w von M mit der Eigenschaft $w \leqq z$ für jede obere Schranke z von M.

wird und für jedes $x \in L$ die Beziehung $x = f\,p(L_x)$ gilt. Jedes Anfangsstück H einer Kette L ist selbst eine Kette; denn für $x \in H$ gilt $H_x = L_x$.

Jetzt seien L und M Ketten und A ein Anfangsstück von L mit der Eigenschaft, daß jedes echte Anfangsstück von A auch Anfangsstück von M ist. Hat A kein letztes Element, gibt es also zu jedem $x \in A$ ein $y \in A$ mit $x < y$, so gehört jedes $x \in A$ einem echten Anfangsstück von A — nämlich A_y — an: A ist Vereinigung echter Anfangsstücke. Diese sind aber Anfangsstücke von M, und daher ist auch A eines. Hat aber A ein letztes Element y, ist — in verständlicher Bezeichnung — $A = \{A', y\}$, so ist A', als echtes Anfangsstück von A, Anfangsstück von M. Ist nun $M \neq A'$, also $M = \{A', z, \ldots\}$, so gilt $y = f\,p(L_y) = f\,p(A') = f\,p(M_z) = z$, d. h. A ist Anfangsstück von M.

Die Anfangsstücke von L sind die Abschnitte L_x mit $x \in L$ und L selbst. Aus $x < y$ folgt $L_x \subset L_y \subset L$; d. h. die Anfangsstücke von L sind durch die Beziehung $\subset$ wohlgeordnet. Ist irgend ein Anfangsstück von L nicht Anfangsstück von M, so gibt es ein erstes A mit dieser Eigenschaft. Seine echten Anfangsstücke sind Anfangsstücke von M. Nach dem vorigen Absatz kann dies nur dann stattfinden, wenn $M = A'$ und daher M Anfangsstück von L ist. Dies tritt sicher ein, wenn L nicht Anfangsstück von M ist; denn L ist ja selbst Anfangsstück von L. Also ist von irgend zwei Ketten immer eine Anfangsstück der anderen.

Wir zeigen jetzt, daß die Vereinigung K aller Ketten selbst eine Kette ist. Zunächst ist K vollständig geordnet. Von zwei Elementen $x \in K$ und $y \in K$ gehört nämlich jedes einer Kette an. Von diesen enthält eine die andere und daher sowohl x wie y; da sie als Kette vollständig geordnet ist, gilt eine der Beziehungen $x \leqq y$ und $y < x$. Sei nun H eine nicht leere Teilmenge von K, etwa $h \in H$. Dann gehört h einer Kette L an. Der Durchschnitt $H \cap L$ hat als nicht leerer Teil der Kette L ein erstes Element h_0. Ist nun $h_1 \in H$, so ist entweder $h_1 \in L$, $h_1 \in H \cap L$ und daher $h_0 \leqq h_1$, oder h_1 gehört einer Kette M an, die nicht Teil von L ist. Dann ist L Anfangsstück von M und daher $h_0 < h_1$, wegen $h_0 \in L$, $h_1 \bar{\in} L$ [6]). Damit ist gezeigt, daß K wohlgeordnet ist. Ist $x \in K$, so gehört x einer Kette $L \subset K$ an. Ist $y \in K$, $y \bar{\in} L$, so gehört y einer Kette M an, die nicht Teil von L ist. Also ist L Anfangsstück von M und daher $x < y$. Ist also $u \in K$, $u \leqq x$, so ist $u \in L$. Da $x \in K$ beliebig war, ist L Anfangsstück von K. Daraus folgt $K_x = L_x$, $x = f\,p(L_x) = f\,p(K_x)$. Damit ist nachgewiesen, daß K eine Kette ist.

Nun ist $w = f\,p(K)$ wegen $p(K) \leqq f\,p(K)$ obere Schranke von K. Wäre $w \bar{\in} K$, so wäre $\{K, w\}$ eine Kette, im Widerspruch zu der Eigen-

[6]) Das in der Druckerei augenblicklich nicht vorhandene durchstrichene $\in$ mit der Bedeutung „ist nicht Element von" wird hier durch das überstrichene ersetzt.

schaft von K, Vereinigung aller Ketten zu sein. Also ist $w \in K$ und daher $w \leqq p(K) \leqq f p(K) = w$. Daraus folgt $w = p(K) = f p(K) = f w$. Damit ist der Satz C bewiesen.

Den hier gegebenen Beweis kann man in einen Beweis des Wohlordnungssatzes verwandeln. Man braucht nur jeder echten Teilmenge M der gegebenen Menge E ein Element $g(M) \in E - M$ zuzuordnen und den Begriff der Kette folgendermaßen abzuändern: Kette heißt jede Menge $L \subseteq E$, der eine Wohlordnung aufgeprägt ist derart, daß für jedes $x \in L$ die Beziehung $x = g(L_x)$ gilt. Dann gehen alle Schlüsse wie vorher. Gegenüber den vorliegenden Darstellungen des ersten Zermeloschen Beweises hat man dabei einen Vorteil: man braucht nicht vorher zu beweisen, daß von zwei wohlgeordneten Mengen eine einem Anfangsstück der anderen ähnlich ist.

(Eingegangen am 10. Oktober 1949.)

Mathematische Zeitschrift, Band 53, Heft 2, S. 114—121 (1950).

Über die metrische Struktur eines Kontinuums endlicher Länge im Euklidischen Raum.

Herrn Erich Kamke zum 60. Geburtstag am 18. August 1950 gewidmet.

Von

Georg Nöbeling in Erlangen.

Es sei K ein Kontinuum[1]) endlicher Länge $L(K)$[2]) im n-dimensionalen Euklidischen Raum R_n. Eine (ungerichtete) Gerade $T(p)$ durch einen Punkt p von K nennen wir *direkte Tangente*[3]) von K in p, wenn jede der beiden in $T(p)$ enthaltenen Halbgeraden H mit dem Anfangspunkt p, jedoch keine weitere Halbgerade H mit dem Anfangspunkt p die folgende Eigenschaft besitzt: Es existiert in K eine Folge $\{p_i\}$ $(i = 1, 2, \ldots)$ von Punkten $p_i \neq p$ derart, daß $\lim p_i = p$ gilt und die Folge $\{H_i\}$ $(i = 1, 2, \ldots)$ der Halbgeraden H_i durch p_i mit dem Anfangspunkt p gegen H konvergiert. Wir behaupten nun erstens folgenden

Satz 1. *Die Menge A aller Punkte p von K, in denen K keine direkte Tangente $T(p)$ hat, besitzt das lineare* Carathéodory*-Maß*

---

[1]) Ein Kontinuum K des R_n ist eine zusammenhängende, beschränkte und abgeschlossene Punktmenge des R_n; wie viele Autoren, verlangen wir außerdem, daß K mehr als einen Punkt enthält.

[2]) Für ein Kontinuum K des R_n stimmen die bekannten linearen Maße für Punktmengen überein (G. Nöbeling, Jahresber. d. Deutschen Mathematikervereinigung 52, 1942, S. 132—160); dieser gemeinsame Wert $L(K)$ ist die Länge von K. Es genügt für unsere Zwecke, für L das lineare Maß L_c von C. Carathéodory (Göttinger Nachrichten 1914, S. 404) zu nehmen, das man auf folgende zwei Arten definieren kann: 1. Bezeichnen wir für jedes $\varrho > 0$ mit d_ϱ die untere Grenze der Durchmessersummen $\Sigma d(M_i)$ für alle Darstellungen $M = M_1 + M_2 + \cdots$ der gegebenen Punktmenge M als Vereinigung von endlich oder abzählbar vielen Teilmengen M_i mit Durchmessern $d(M_i) < \varrho$, so ist $L_c(M) = \lim_{\varrho \to 0} d_\varrho$. 2. Bezeichnen wir für jedes $\varrho > 0$ mit d'_ϱ die untere Grenze der Durchmessersummen $\Sigma d(O_i)$ für alle Überdeckungen $O_1 + O_2 + \cdots > M$ von M durch endlich oder abzählbar viele konvexe, offene Mengen O_i mit Durchmessern $d(O_i) < \varrho$, so ist $L_c(M) = \lim_{\varrho \to 0} d'_\varrho$. (Die Äquivalenz dieser beiden Definitionen folgt unmittelbar daraus, daß jede Menge denselben Durchmesser hat wie ihre konvexe Hülle.)

[3]) „direkt" im Sinne der direkten Differentialgeometrie.

$L_C(A) = 0$[4]). *Für jedes* $\varepsilon > 0$ *existiert eine perfekte Teilmenge* P *von* $K - A$ *derart, daß* $L_C(K - P) < \varepsilon$ *und* $T(p)$ *auf* P *stetig ist*[5])[6]).

Zweitens behaupten wir für jedes Kontinuum K endlicher Länge $L(K)$ im R_n:

Satz 2. *Die Menge* B *aller End- und Verzweigungspunkte*[7]) *von* K *hat das Maß* $L_C(B) = 0$[8]).

Beweis des Satzes 1.

Wir wählen im R_n ein rechtwinkliges, CARTESISCHES Koordinatensystem $x_1, \ldots, x_n$.

Wegen $L(K) < +\infty$ ist K auch eine stetige Kurve endlicher Durchlaufungslänge, d. h. es existiert eine eindeutige stetige Abbildung $f(s) = [x_\nu = f_\nu(s);\ \nu = 1, \ldots, n]$ eines Intervalls $I = [0 \leqq s \leqq S]$ auf K derart, daß die zugehörige (Durchlaufungs-)Länge

$$L_f(K) = \text{ob. Gr.} \sum_{\varrho=1}^{\varrho_0} \sqrt{\sum_{\nu=1}^{n} (f_\nu(s_\varrho) - f_\nu(s_{\varrho-1}))^2} \quad (s_0 = 0 < s_1 < \cdots < s_{\varrho_0} = S)$$

endlich ist[9]). Dabei können wir sofort annehmen, daß der Parameter s die Länge in $K = f(I)$ ist, d. h. daß für jeden Parameterwert s' aus I die Länge $L_f(K')$ der Kurve $K' = f(I')$, $I' = [0 \leqq s \leqq s']$ gleich s' ist. Dann existiert bekanntlich[10]) eine Teilmenge I_0 von I mit dem LEBESGUE-Inhalt $m I_0 = 0$ derart, daß für jedes s aus $I - I_0$ die Ableitungen $f'_\nu(s)$

[4]) Für $n = 2$ ist diese Behauptung enthalten in einem allgemeineren Satz von A. S. BESICOVITCH (Math. Ann. 115, (1938, 296—329, Satz 17). Für ein Kontinuum (der Ebene) ist der Tangentenbegriff von BESICOVITCH mit unserem Begriff der direkten Tangente äquivalent.

[5]) Dieser Satz ist das „direkte" Analogon zu dem bekannten Satz über die Existenz und Stetigkeit der Tangenten bei eindeutigen stetigen Streckenbildern endlicher Länge (vgl. z. B. O. HAUPT und G. AUMANN, Differential- und Integralrechnung, Bd. II, 1. Auflage, Berlin 1932, S. 98).

[6]) Es ist $L(A) = 0$ auch für jedes a. a. O. [2]), § 1, I. 2 und III—IX angegebene lineare Maß L. Denn nach der zweiten Behauptung des Satzes 1 existiert für jedes natürliche n eine perfekte Teilmenge P_n von $K - A$ mit $L_C(K - P_n) < \frac{1}{n}$. Für die Vereinigung Q der P_n gilt dann $L_C(K - Q) = 0$. Da $L(K) < +\infty$ und $K - Q$ eine BORELsche Menge ist, gilt $L_C(K - Q) = L(K - Q)$ (a. a. O. [2]), Satz 4), also $L(K - Q) = 0$. Wegen $A < K - Q$ ist also $L(A) = 0$. Analog ist $L(K - P) < \varepsilon$.

[7]) Im Sinne der Kurventheorie; vgl. S. 118 und K. MENGER, Kurventheorie, Leipzig-Berlin 1932, S. 96 ff.

[8]) Es ist $L(B) = 0$ auch für jedes der a. a. O. [2]), § 1, I. 2 und III—IX angegebenen Maße L. Denn B ist eine BORELsche Menge (K. MENGER, a. a. O. [7]), S. 104). Also ist $L_C(B) = L(B)$ (a. a. O. [2]), Satz 4).

[9]) S. BOCHNER, Jahresber. d. Deutschen Mathematikervereinigung 40, 1931, S. 245—249.

[10]) a. a. O. [5]).

($\nu = 1, \ldots, n$) existieren,

$$\sum_{\nu=1}^{n} (f'_\nu(s))^2 = 1 \tag{1}$$

ist und daher im Punkt $f(s)$ die Tangente $T(s)$ (im üblichen Sinne) an die Kurve $f(I)$ existiert (d. i. die gerichtete Gerade durch den Punkt $f(s)$ mit den Richtungscosinus $f'_1(s), \ldots, f'_n(s)$). Wir nennen diese Tangente $T(s)$ (zwecks Unterscheidung von der direkten Tangente) die *Durchlaufungstangente* der Kurve $f(I) = K$ an der Stelle s, weil sie mit Hilfe einer Durchlaufung von K, d. h. einer Abbildung f von I auf K definiert ist [11]).

Es sei $f(I_0) = K_0$ gesetzt. Dann ist

$$L_C(K_0) = 0. \tag{2}$$

Es sei nämlich ein $\varepsilon > 0$ beliebig gegeben. Wegen $m I_0 = 0$ existiert eine Überdeckung von I_0 durch höchstens abzählbar viele Intervalle $I_1, I_2, \ldots \subset I$ mit Längen $L(I_i)$, deren Summe $< \varepsilon$ ist. Da die Abbildung $f(s)$ dehnungslos ist, besitzen die Bildmengen $f(I_i) = M_i$ Durchmesser $d(M_i) \leqq L(I_i)$; die Summe dieser Durchmesser ist also $< \varepsilon$. Die Mengen M_i bilden eine Überdeckung von K_0. Also gilt (2).

Nun sei K^k die Menge aller Punkte p aus K mit der Eigenschaft, daß es in I mindestens k Parameterwerte s gibt mit $f(s) = p$ ($k = 1, 2, \ldots, +\infty$). Wir behaupten:

$$L_C(K^\infty) = 0. \tag{3}$$

Für jeden Punkt p von K^∞ ist die Urbildmenge $f^{-1}(p)$ unendlich und abgeschlossen. Die (abgeschlossene) Menge der Häufungspunkte von $f^{-1}(p)$ ist also nicht leer und eine Teilmenge von $f^{-1}(p)$. Ist daher $s_0(p)$ ein Häufungswert von $f^{-1}(p)$, etwa der kleinste Häufungswert, so gilt $f(s_0(p)) = p$ und es existiert in $f^{-1}(p)$ eine Folge $s_i(p)$ ($i = 1, 2, \ldots$) mit $s_i(p) \neq s_0(p)$ und $\lim s_i(p) = s_0(p)$. Da $f(s_i(p)) = p = f(s_0(p))$ gilt für jedes i, ist $f'_1(s_0(p)) = \cdots = f'_n(s_0(p)) = 0$, falls diese Ableitungen existieren. Mithin ist $s_0(p)$ in der Menge I_0 enthalten, da für jedes s aus $I - I_0$ die Ableitungen $f'_\nu(s)$ existieren und ihre Quadratsumme gleich 1 ist. Die Menge I^∞ der $s_0(p)$, wobei p die Menge K^∞ durchläuft, ist also eine Teilmenge von I_0, wegen $m I_0 = 0$ mithin eine Nullmenge. Nun ist $f(I^\infty) = K^\infty$ und die Abbildung f ist dehnungslos. Wie I^∞ kann also auch K^∞ für jedes $\varepsilon > 0$ dargestellt werden als

[11]) Der Zusammenhang zwischen der direkten Tangente und der Durchlaufungstangente ist folgender. Existiert in einem Punkt p von K die direkte Tangente $T(p)$ und existiert an einer Stelle $s \in I$ mit $f(s) = p$ die Durchlaufungstangente $T(s)$, so fällt $T(s)$ in $T(p)$ (d. h. die durch Weglassung der Orientierung aus $T(s)$ entstehende Gerade $|T(s)|$ ist gleich $T(p)$). Existiert an jeder Stelle $s \in I$ mit $f(s) = p$ die Durchlaufungstangente $T(s)$ und fallen alle diese $T(s)$ in eine und dieselbe Gerade T, so ist T direkte Tangente von K in p.

Vereinigung höchstens abzählbar vieler Mengen, deren Durchmessersumme $< \varepsilon$ ist. Folglich gilt (3).

Wegen

$$A = A \,.\, K = A \,.\, \left(\sum_{k=1}^{\infty} (K^k - K^{k+1}) + K^\infty \right)$$

$$< \sum_{k=1}^{\infty} A \,.\, (K^k - K^{k+1}) + K^\infty < \sum_{k=1}^{\infty} (A \,.\, (K^k - K^{k+1}) - K_0) + K_0 + K^\infty$$

ist

$$L_C(A) \leqq \sum_{k=1}^{\infty} L_C(A \,.\, (K^k - K^{k+1}) - K_0) + L_C(K_0) + L_C(K^\infty).$$

Wegen (2) und (3) genügt es also zum Beweis der ersten Behauptung $L_C(A) = 0$ des Satzes 1 zu beweisen, daß für jedes natürliche k

$$L_C(A \,.\, (K^k - K^{k+1}) - K_0) = 0$$

ist. Und hierfür wiederum genügt es zu zeigen:

$$A \,.\, (K^k - K^{k+1}) - K_0$$

ist höchstens abzählbar.

Angenommen, für ein natürliches $k = k_0$ wäre die Menge

$$A \,.\, (K^{k_0} - K^{k_0+1}) - K_0 = J \tag{4}$$

unabzählbar.

Ist p ein beliebiger Punkt von J, so hat p wegen $J < K^{k_0} - K^{k_0+1}$ in I genau k_0 Urbilder $s_1(p), \ldots, s_{k_0}(p)$; wir können sie sofort wachsend nummeriert annehmen. Für jedes natürliche j bezeichnen wir mit J^j die Menge aller Punkte von J mit der Eigenschaft, daß

$$s_{i+1}(p) - s_i(p) \geqq \frac{1}{j} \qquad (i = 1, \ldots, k_0 - 1) \tag{5}$$

ist. Dann ist $J^1 + J^2 + \cdots = J$. Nach (4) ist also für mindestens ein $j = j_0$ die Menge J^{j_0} unabzählbar. Dann enthält J^{j_0} einen Punkt p, welcher der Limes einer Folge $\{p_m\}$ $(m = 1, 2, \ldots)$ von Punkten $p_m \neq p$ aus J^{j_0} ist. Aus $p_m \neq p$ folgt

$$s_i(p_m) \neq s_i(p). \tag{6}$$

Da sowohl für den Punkt p, als auch für die Punkte p_m die Ungleichungen (5) gelten, folgt aus $\lim p_m = p$

$$\lim s_i(p_m) = s_i(p). \tag{7}$$

Da J zu $K_0 = f(I_0)$ fremd ist, liegt keines der Urbilder $s_1(p), \ldots, s_{k_0}(p)$ in I_0. Folglich existieren die Durchlaufungstangenten $T(s_1(p)), \ldots, T(s_{k_0}(p))$ an die Kurve $f(I) = K$ an den Stellen $s_1(p), \ldots, s_{k_0}(p)$. Diese Tangenten fallen nicht sämtlich in eine Gerade T; denn dann wäre diese Gerade direkte Tangente an das Kontinuum K in p, im Wider-

spruch dazu, daß p als Punkt von J in A liegt. Es mögen etwa die Tangenten $T(s_1(p))$ und $T(s_2(p))$ nicht in eine und dieselbe Gerade fallen, d. h. sie mögen sich (in p) schneiden. Dann existieren zwei im Intervall I offene Intervalle I_1 und I_2 derart, daß $s_1(p)$ in I_1, $s_2(p)$ in I_2 liegt und aus $s' \in I_1$, $s'' \in I_2$, $f(s') = f(s'')$ folgt $s' = s_1(p)$ und $s'' = s_2(p)$. Für hinreichend großes m gilt nun aber $s_1(p_m) \in I_1$ und $s_2(p_m) \in I_2$ wegen (7) und es ist $f(s_1(p_m)) = p_m = f(s_2(p_m))$. Also ist $s_1(p_m) = s_1(p)$ und $s_2(p_m) = s_2(p)$. Dies widerspricht aber (6). Damit ist die Behauptung $L_C(A) = 0$ des Satzes 1 bewiesen.

Die Funktionen $f_\nu(s)$ sind von beschränkter Variation wegen $|f_\nu(s') - f_\nu(s'')| \leqq |s' - s''|$, also darstellbar als Differenzen monotoner (beschränkter) Funktionen. Folglich sind die Ableitungen $f'_\nu(s)$ in $I - I_0$ meßbar[12]). Die Menge $I - I_0$ enthält daher eine perfekte Teilmenge I_1 derart, daß die Menge $I_2 = I - I_1$ ein Maß $mI_2 < \frac{\varepsilon}{2}$ hat und die Ableitungen $f'_\nu(s)$ auf I_1 stetig sind[13]). Da die Abbildung $f(s)$ dehnungslos ist, hat das Bild $K_2 = f(I_2)$ von I_2 ein Maß $L_C(K_2) < \frac{\varepsilon}{2}$. Wegen $L_C(A) = 0$ ist dann auch $L_C(A + K_2) < \frac{\varepsilon}{2}$. Da auf Grund der zweiten Definition [4]) des Carathéodory-Maßes die Menge $A + K_2$ eine maßgleiche Obermenge $< K$ besitzt, welche ein G_δ ist, hat $A + K_2$ auch eine Obermenge $O < K$, welche in K offen ist und für welche $L_C(O) < \varepsilon$ gilt. Wegen $K - O < K - A$ ist in jedem Punkt p von $K - O$ die direkte Tangente $T(p)$ an K vorhanden; wegen $K - O < K - K_2 < f(I_1)$ fällt für jedes $s \in I_1$ mit $f(s) = p$ die Durchlaufungstangente $T(s)$ in die direkte Tangente $T(p)$; da $T(s)$ auf I_1 stetig ist, so ist also $T(p)$ auf $K - O$ stetig. Die Menge $K - O$ ist in K abgeschlossen und enthält eine maßgleiche perfekte Teilmenge P; für diese trifft die zweite Behauptung des Satzes 1 zu.

Beweis des Satzes 2.

Aus $L(K) < +\infty$ folgt, daß das Kontinuum K eine reguläre Kurve ist[2]). Das heißt folgendes. Ist p ein beliebiger Punkt von K, so existieren Umgebungen von p mit beliebig kleinen Durchmessern, deren Begrenzungen mit K je nur endlich viele Punkte gemein haben. Existieren natürliche Zahlen ν derart, daß p Umgebungen mit beliebig kleinen Durchmessern hat, deren Begrenzungen mit K höchstens ν Punkte gemein haben, so heißt die kleinste Zahl ν mit dieser Eigenschaft die (Verzweigungs-) Ordnung von p; existieren solche Zahlen nicht, so heißt p von der Ordnung ω. Die Punkte der Ordnung 1 werden Endpunkte, die Punkte mit Ordnungen > 2 oder $= \omega$ Verzweigungspunkte

[12]) C. Carathéodory, Reelle Funktionen, Leipzig und Berlin 1918, S. 773.

[13]) a. a. O. [12]), S. 410.

von K genannt [7]). Wir bezeichnen die Menge aller Endpunkte von K mit K_e und die Menge aller Verzweigungspunkte von K mit K_v. Die Behauptung des Satzes 2 lautet: $L_C(K_e + K_v) = 0$. Es genügt, die beiden Gleichungen

$$L_C(K_e) = 0 \tag{8}$$

und

$$L_C(K_v) = 0 \tag{9}$$

zu beweisen.

Es sei p ein Punkt von K, in welchem K eine direkte Tangente $T(p)$ besitzt. Ist dann $E(p)$ die zu $T(p)$ senkrechte Hyperebene durch p und $U(p, r)$ die offene Vollkugel mit dem Mittelpunkt p und dem Radius r, so hat für hinreichend kleines r die Hyperkreisscheibe $H(p,r) = U(p,r) \,.\, E(p)$ mit K nur den Punkt p gemein, wie aus der Definition der direkten Tangente $T(p)$ folgt. Wählen wir auf beiden Seiten der Hyperebene $E(p)$ in hinreichender Nähe von p je einen Punkt p^- bzw. p^+ von K (sie existieren nach der Definition der direkten Tangente $T(p)$), so können wir sie innerhalb K durch einen Bogen B verbinden, der eine Teilmenge von $U(p, r)$ ist; denn K ist wegen $L(K) < +\infty$ im Kleinen zusammenhängend [2]). Da die Hyperebene $E(p)$ die Punkte p^- und p^+ trennt, ist der Durchschnitt $B \,.\, E(p)$ nicht leer, wegen $B \subset U(p, r)$ also auch der Durchschnitt $B \,.\, H(p, r)$ nicht leer. Da aber $K \,.\, H(p, r) = (p)$ ist, so folgt $B \,.\, H(p, r) = (p)$. Der Punkt p liegt also auf dem Bogen B. Er ist aber von den beiden Endpunkten p^- und p^+ von B verschieden. Wegen $B \subset K$ ist also p kein Endpunkt von K. Dies gilt für jeden Punkt p von K mit direkter Tangente $T(p)$, also für jeden Punkt von $K - A$. Die Menge K_e aller Endpunkte von K ist daher eine Teilmenge der Menge A aller Punkte p von K, in denen K keine direkte Tangente $T(p)$ hat. Da $L_C(A) = 0$ ist nach dem Satz 1, folgt (8).

Ist p ein Punkt von K, in welchem K eine direkte Tangente $T(p)$ hat, so bedeute $F(p)$ die Menge, bestehend aus dem Punkt p und allen Punkten $q \neq p$ des Raumes, für welche der (kleinere) Winkel zwischen $T(p)$ und der Verbindungsstrecke von p und q mindestens gleich $\frac{\pi}{8}$ ist. Für jedes hinreichend kleine r ist der Durchschnitt der Menge

$$J(p, r) = U(p, r) \,.\, F(p)$$

mit K die einpunktige Menge (p), wie aus der Definition von $T(p)$ folgt.

Für jedes natürliche m sei nun K_m die Menge aller Punkte p von K, in denen K eine direkte Tangente $T(p)$ hat und für welche

$$K \,.\, J\left(p, \frac{1}{m}\right) = (p) \tag{10}$$

ist. Dann ist $K - A = K_1 + K_2 + \cdots$, also $K = A + \sum K_m$, folglich $K_v = K_v \,.\, A + \sum K_v \,.\, K_m \subset A + \sum K_v \,.\, K_m$ und daher $L_C(K_v) \leqq L_C(A) + \sum L_C(K_v \,.\, K_m)$. Da $L_C(A) = 0$ ist nach Satz 1, genügt es zum Beweis

von (9) zu zeigen, daß $L_C(K_\nu . K_m) = 0$ ist für jedes $m = 1, 2, \ldots$. Wir nehmen an, es gäbe eine natürliche Zahl m, für welche

$$L_C(K_\nu . K_m) > 0 \tag{11}$$

ist, und leiten daraus einen Widerspruch her. Damit ist dann die Gleichung (9) bewiesen.

Nun ist

$$L(K) = \text{ob. Gr.} \sum_{j=1}^{k} L(B_j)$$

wobei $\sum L(B_j)$ die Längensummen aller Systeme endlich vieler, paarweise fremder Teilbogen B_j $(j = 1, \ldots, k)$ von K durchläuft [2]). Mithin können wir endlich viele Bogen B_j $(j = 1, \ldots, k)$ so wählen, daß

$$L_C(K - \sum B_j) = L_C(K) - L_C(\sum B_j) = L(K) - \sum L(B_j) < L_C(K_\nu . K_m)$$

und folglich auch

$$L_C(K_\nu . K_m - \sum B_j) < L_C(K_\nu . K_m)$$

ist. Dann ist

$$L_C(K_\nu . K_m . \sum B_j) > 0$$

und folglich

$$L_C(K_\nu . K_m . B_j) > 0$$

für mindestens ein j. Wir schreiben B statt B_j. Dann ist also B ein Teilbogen von K mit der Eigenschaft

$$L_C(K_\nu . K_m . B) > 0.$$

Nach dem Satz 1 existiert eine perfekte Teilmenge P von $K - A$, auf welcher überall die direkte Tangente $T(p)$ an K existiert und stetig ist, derart, daß $L_C(K - P) < L_C(K_\nu . K_m)$, also

$$L_C(K_\nu . K_m . B . P) > 0 \tag{12}$$

ist. Nach dem Satz 1, jetzt angewandt auf B statt K, gilt für die Menge D aller Punkte von p, in denen keine direkte Tangente an B existiert, die Gleichung $L_C(D) = 0$, also wegen (12)

$$L_C(K_\nu . K_m . (B - D) . P) > 0.$$

Wegen dieser Ungleichung ist die Teilmenge $K_\nu . K_m . (B - D) . P$ des Bogens B unabzählbar. Folglich enthält diese Menge einen Punkt p_0 und für jedes natürliche i zwei Punkte p_i^- und p_i^+ derart, daß gilt:

$$\lim p_i^- = p_0 = \lim p_i^+, \tag{13}$$

$$p_0 \text{ trennt } p_i^- \text{ und } p_i^+ \text{ in } B \qquad (i = 1, 2, \ldots). \tag{14}$$

Da wegen $B \subset K$ die direkte Tangente von B in p_0 gleich der direkten Tangente von K in p_0 ist und die Hyperebene $E(p)$ durch p_0 diese

Tangente (senkrecht) schneidet, so folgt aus (14):

(15) $E(p_0)$ trennt p_i^- und p_i^+ im Raum $(i > i_0)$,

falls i_0 hinreichend groß ist.

Wegen der Stetigkeit der direkten Tangente $T(p)$ an K auf der Menge P gilt wegen (13), $p_i^- \in P$, $p_0 \in P$ und $p_i^+ \in P$

$$\lim T(p_i^-) = T(p_0) = \lim T(p_i^+). \tag{16}$$

Es sei nun ein $\varepsilon > 0$ beliebig gegeben. Wir verbinden für jedes i die Punkte p_i^- und p_i^+ mit jedem Punkt der Hyperebene $E(p_0)$, der von p_0 den Abstand $\frac{\varepsilon}{2}$ hat, durch eine abgeschlossene Strecke und bezeichnen die Vereinigungsmenge dieser Strecken mit U_i^*. Wegen (15) ist U_i^* die Begrenzung einer Umgebung U_i von p_0. Nach (13) ist bei hinreichend großem i_0 für $i > i_0$ der Durchmesser von U_i gleich ε. Weiter ist bei hinreichend großem i_0 für $i > i_0$ die Menge U_i^* wegen (13), (15) und (16) eine Teilmenge der Menge $J\left(p_i^-, \frac{1}{m}\right) + J\left(p_i^+, \frac{1}{m}\right)$. Nach (10) ist also $K \,.\, U_i^* = (p_i^-) + (p_i^+)$. Damit haben wir gezeigt: Es existiert eine Umgebung von p_0 mit einem Durchmesser $\leqq \varepsilon$, deren Begrenzung mit K genau zwei Punkte gemein hat; dabei ist ε eine beliebige positive Zahl. Also hat p_0 eine Verzweigungsordnung $\leqq 2$. Dies steht aber im Widerspruch dazu, daß p_0 ein Punkt von K_v, also ein Verzweigungspunkt von K ist. Damit ist der gesuchte Widerspruch hergeleitet und die Gleichung (9) bewiesen.

(Eingegangen am 22. Oktober 1949.)

Mathematische Zeitschrift, Band 53, Heft 2, S. 122—130 (1950).

Wronskische Determinante und lineare Abhängigkeit.

Herrn Erich Kamke zum 60. Geburtstag am 18. August 1950 gewidmet.

Von

Otto Haupt in Erlangen.

1. Vorbemerkung.

Es sei $\mathfrak{J}$ eine zusammenhängende Teilmenge der reellen Zahlgeraden $\mathfrak{E}_1$. Ferner seien $f_\nu(x)$, $\nu = 1, \ldots, n$; $n \geqq 1$, eindeutige, reelle, $(n-1)$-mal eigentlich differenzierbare Funktionen von x in $\mathfrak{J}$. Bekanntlich läßt sich dann aus der linearen Abhängigkeit [1]) der f_ν auf das in $\mathfrak{J}$ identische Verschwinden der Wronskischen Determinante $W_n(x)$ der $f_1, \ldots, f_n$ schließen, während umgekehrt [2]) aus $W_n(x) \equiv 0$ nur die Existenz von (abzählbar vielen) zusammenhängenden Teilen $\mathfrak{J}'$ von $\mathfrak{J}$ folgt, auf deren jedem die f_ν linear abhängig sind [3]). Es gibt nun im allgemeinen verschiedene Systeme solcher $\mathfrak{J}'$. Indes läßt sich, wie wir zeigen werden, mit Hilfe der Nullstellenmengen von $W_n(x)$ und den Wronskischen Unterdeterminanten von $W_n(x)$ eindeutig ein System solcher $\mathfrak{J}'$ auf sehr einfache Weise bestimmen; hierbei ergibt sich zugleich wieder ein Beweis [4]) der erwähnten Umkehrung. Bei diesen Betrachtungen werden nur sehr wenige Eigenschaften des Differentialquotienten bzw. der differenzierbaren Funktionen benutzt. Die Heraushebung dieser Eigenschaften führt zu einer Verallgemeinerung der vorerwähnten Sätze. Dies möge nachstehend dargelegt werden.

2. Voraussetzungen. Bezeichnungen.

2. 1. Voraussetzungen. Es seien $\mathfrak{R}$ und $\mathfrak{K}$ (nicht leere) Gesamtheiten von Elementen; ferner sei $\mathfrak{k}$ ein System von Teilmengen $\mathfrak{T}$ von $\mathfrak{R}$. Weiter sei $\mathfrak{a}$ das System aller eindeutigen Abbildungen (Funktionen) [5])

---

[1]) Vgl. die Definition in Nr. 2. 3. des Textes.

[2]) M. Fréchet, Sur la portée exacte de l'évanouissement d'un Wronskien dans un intervalle, Bull. Calcutta math. Soc. 29 (1937), 121—124.

[3]) Die $\mathfrak{J}'$ können so gewählt werden, daß ihre Vereinigung in $\mathfrak{J}$ dicht liegt.

[4]) Dieser Beweis verläuft etwas anders als der von Herrn Fréchet a. a. O.[2]) angegebene. Vgl. im Text Nr. 3., Satz 3.

[5]) Es bezeichnet $f(X) \mid \mathfrak{T}$ eine Funktion mit dem Definitionsbereich $\mathfrak{T}$.

$f(X) \mid \mathfrak{T}$ von $\mathfrak{T}$ in $\mathfrak{R}$ für alle $\mathfrak{T} \in \mathfrak{t}$. Es sollen nun $\mathfrak{R}$, $\mathfrak{t}$, $\mathfrak{K}$ und $\mathfrak{a}$ den folgenden Axiomen genügen:

($\mathbf{A_1}$) Es ist $\mathfrak{R}$ ein topologischer Raum (ohne Trennungsaxiome)[6], der *in jedem seiner Punkte* stark zusammenhängend im Kleinen[7] ist, d. h. jeder Punkt von $\mathfrak{R}$ besitzt beliebig kleine[8] *zusammenhängende* Umgebungen. Weiter sei $\mathfrak{t}$ der σ-Ring der in $\mathfrak{R}$ offenen Mengen. Schließlich sei $\mathfrak{K}$ ein topologischer Körper (im Sinne der Algebra)[9].

($\mathbf{A_2}$) Es existiert ein Teilsystem $\mathfrak{d}$ von $\mathfrak{a}$ mit folgenden Eigenschaften:

(I) mit $f(X) \mid \mathfrak{T}$ gehört auch jede ihrer „Verengerungen" $f(X) \mid \mathfrak{T}''$ zu $\mathfrak{d}$, für welche $\mathfrak{T}'' \in \mathfrak{t}$ (und $\mathfrak{T}'' \subset \mathfrak{T}$) ist;

(II) Jedes zu $\mathfrak{d}$ gehörige $f(X) \mid \mathfrak{T}$ ist stetig[10] in $\mathfrak{T}$;

(III) Es ist $\mathfrak{d}$ rational geschlossen, d. h. mit $f_1 \in \mathfrak{d}$, $f_2 \in \mathfrak{d}$ ist auch $f_1 + f_2 \in \mathfrak{d}$, $f_1 \cdot f_2 \in \mathfrak{d}$ und $f_1 : f_2 \in \mathfrak{d}$. Dabei ist als Definitionsbereich zu nehmen: Für $f_1 + f_2$ und $f_1 \cdot f_2$ der (in $\mathfrak{R}$ offene) Durchschnitt $\mathfrak{T}'$ der Definitionsbereiche von f_1 und f_2, hingegen für $f_1 : f_2$ die (in $\mathfrak{R}$ offene) Menge $\mathfrak{T}_0'$ aller $X \in \mathfrak{T}'$ mit $f_2(X) \neq 0$.

($\mathbf{A_3}$) Es existiert eine eindeutige Abbildung $D(f) = D(X; f)$ von $\mathfrak{d}$ in $\mathfrak{a}$, etwa von $\mathfrak{d}$ auf $\mathfrak{d}' \subset \mathfrak{a}$ (m. a. W. eine Operation, vermöge deren $f(X) \mid \mathfrak{T} \in \mathfrak{d}$ in $D(X; f) \mid \mathfrak{T} \in \mathfrak{d}'$ übergeht) mit folgenden Eigenschaften:

(I) Es ist $\mathfrak{d}'$ ganz-rational geschlossen (also abgesehen von der Division, vgl. (A_2) (III)); dabei gilt noch

$$D(f_1 + f_2) = D(f_1) + D(f_2),\ D(f_1 f_2) = f_2 D(f_1) + f_1 D(f_2).$$

---

[6]) sogen. t-Raum; vgl. Haupt-Aumann-Pauc, Diff. u. Integr. Rechn. 2. Aufl., I. Bd., Nr. 6. 2. 1. (1948). Im Text wird der *offene Kern* bzw. die *abgeschlossene Hülle* von $\mathfrak{M} \subset \mathfrak{R}$ mit $\underline{\mathfrak{M}}$ bzw. mit $\overline{\mathfrak{M}}$ bezeichnet und mit $\mathfrak{O}$ die *leere Menge*.

[7]) Vgl. (für metrische Räume) K. Menger, Kurventheorie (Leipzig 1932), S. 32. Ist $\mathfrak{R}$ stark zusammenhängend im Kleinen im Punkt $X \in \mathfrak{R}$, so auch *zusammenhängend im Kleinen* in X (d. h. zu jeder Umgebung $\mathfrak{U}$ von $X \in \mathfrak{R}$ existiert eine Umgebung $\mathfrak{V}$ von X derart, daß jeder Punkt von $\mathfrak{V}$ mit X durch eine, in $\mathfrak{U}$ enthaltene zusammenhängende Teilmenge von $\mathfrak{R}$ verbunden ist); die Umkehrung hiervon gilt nicht allgemein (vgl. Menger, a. a. O.).

[8]) d. h. in jeder Umgebung des Punktes X ist eine zusammenhängende Umgebung von X enthalten.

[9]) Unter einem topologischen Körper $\mathfrak{K}$ wird hierbei verstanden ein System von Elementen, dem aufgeprägt sind 1. die Struktur eines kommutativen Körpers im Sinne der Algebra; 2. eine Topologie, bei der stetig sind die Abbildungen $f(X) = -X$, $g(X) = X^{-1}$ von $\mathfrak{K}$ in $\mathfrak{K}$, sowie die Abbildungen $h(X, Y) = X + Y$ und $j(X, Y) = X \cdot Y$ von $\mathfrak{K} \times \mathfrak{K}$ in $\mathfrak{K}$. Vgl. N. Bourbaki, Éléments de mathématique, I. part., Livre III. Topologie générale, Chap. III, § 5. (Als topologisch können insbesondere die bewerteten Körper aufgefaßt werden.)

[10]) Die Definition der Stetigkeit der $f(X)$ bezieht sich auf die Topologie in $\mathfrak{R}$ und in $\mathfrak{K}$ (vgl. Bourbaki, a. a. O.; auch a. a. O.[6]), Nr. 6. 1. ff.)

(II′) Ist $\mathfrak{M}$ offen in $\mathfrak{R}$ sowie zusammenhängend[11]) und[12]) ist $f(X) \equiv$ konst $= c$ in $\mathfrak{M}$, so ist $f \in \mathfrak{d}$ und $D(f) \equiv 0$ in $\mathfrak{M}$.

(II″) ist $\mathfrak{Q}$ *offen* in $\mathfrak{R}$ und zusammenhängend und ist $D(X; f_0) \equiv 0$ in $\mathfrak{Q}$ für ein $f_0 \in \mathfrak{d}$, so ist $f_0(X) \equiv$ konst $= c$ in $\mathfrak{Q}$, wobei $c \in \mathfrak{K}$.

2. 2. Bemerkungen zu Nr. 2. 1.

(1) *Das Axiomensystem* $(A_1) - (A_3)$ *ist widerspruchsfrei.* Ist nämlich $\mathfrak{R}$ der 1-dimensionale euklidische Raum (reelle Zahlgerade) $\mathfrak{E}_1$, ferner $\mathfrak{K}$ der Körper der reellen Zahlen, $\mathfrak{a}$ das System der reellen endlichen (eindeutigen) Funktionen der reellen Veränderlichen x und $D(f) = df/dx$ der Differentialquotient im üblichen Sinne, so sind die Axiome erfüllt, wenn $\mathfrak{d}$ das System der differenzierbaren Funktionen (mit endlichem Differentialquotienten) ist. Weitere Beispiele werden durch die Differentialrechnung in Vektorräumen geliefert[12a]).

(2) Die im Kleinen stark zusammenhängenden Räume $\mathfrak{R}$ (vgl. Nr. 2. 1. (A_1)) lassen sich kennzeichnen wie folgt[13]):

Satz. *Für jeden topologischen Raum* $\mathfrak{R}$ *sind folgende Aussagen gleichwertig:* (**a**) *Es ist* $\mathfrak{R}$ *in jedem Punkt stark zusammenhängend im Kleinen* — (**b**) *Jede Komponente einer jeden (in* $\mathfrak{R}$*) offenen Menge ist selbst offen.*

Anmerkung. Jede offene Teilmenge von $\mathfrak{R}$ ist selbst stark zusammenhängend im Kleinen.

Bew. *Aus* (a) *folgt* (b). Es sei $\mathfrak{M}$ offen und $\mathfrak{Z}$ eine Komponente (d. h. größte zusammenhängende Teilmenge) von $\mathfrak{M}$. Wäre $\mathfrak{Z}$ nicht offen, so existierte ein Punkt $S \in \mathfrak{Z} - \underline{\mathfrak{Z}}$ und es wäre in *jeder* Umgebung $\mathfrak{U}$ von S ein $Y \in \mathfrak{M} - \mathfrak{Z}$ enthalten. Gemäß (a) existiert aber speziell eine *in* $\mathfrak{M}$ *enthaltene*, zusammenhängende Umgebung $\mathfrak{V}$ von S. Wegen der Existenz eines $Y \in \mathfrak{V}(\mathfrak{M} - \mathfrak{Z}) = \mathfrak{V} - \mathfrak{Z}$ wäre $\mathfrak{V} - \mathfrak{Z} \neq \mathfrak{O}$, also $\mathfrak{S} = \mathfrak{Z} + \mathfrak{V}$ eine echte, in $\mathfrak{M}$ enthaltene Obermenge von $\mathfrak{Z}$. Andererseits wäre $\mathfrak{S}$ zusammenhängend, weil $S \in \mathfrak{Z}\mathfrak{V}$, also $\mathfrak{Z}\mathfrak{V} \neq \mathfrak{O}$. Es wäre mithin $\mathfrak{Z}$ nicht Komponente von $\mathfrak{M}$. — *Aus* (b) *folgt* (a). Ist $\mathfrak{U}$ eine beliebige Umgebung von $X \in \mathfrak{R}$ und $\mathfrak{Z}$ die, X enthaltende Komponente von $\mathfrak{U}$, so ist $\mathfrak{Z} < \mathfrak{U}$ offen (gemäß (b)), also $\mathfrak{Z}$ eine (in $\mathfrak{U}$ enthaltene) zusammenhängende Umgebung von X.

(3) Verabredung. Im Hinblick auf Nr. 2. 1., (A_2) (III) soll im folgenden stets vorausgesetzt werden: Funktionen $f_j \mid \mathfrak{T}_j$, die *gleich-*

[11]) Im Spezialfall der reellen Funktionen einer reellen Veränderlichen würde für (II′) die Forderung genügen, daß $\mathfrak{M}$ in sich dicht ist, wenn $D(f) = df/dx$ gesetzt wird.

[12]) Wir sagen, $f(X) \mid \mathfrak{D}$ sei *identisch gleich* c (speziell *verschwinde identisch*, wenn c gleich Null ist) in $\mathfrak{D}$, in *Zeichen* $f(X) \equiv c$, wenn $f(X) = c$ *überall in* $\mathfrak{d}$, d. h. für *jedes* $X \in \mathfrak{D}$.

[12a]) Vgl. H. J. Kowalsky, Differenzenquotienten in lokalkonvexen Vektorräumen, Sitz.-Ber. d. bayer. Akad. d. Wiss., Math.-naturwiss. Kl. 1950.

[13]) Vgl. (für metrische Räume) Menger, a. a. O. [7]), S. 33.

zeitig betrachtet werden, besitzen *gleichen* Definitionsbereich $\mathfrak{T}_j = \mathfrak{T}$; dies bedeutet keine Einschränkung der Allgemeinheit.

2. 3. Bezeichnungen. (1) Wir setzen $D^{(0)}(f) = f$, $D^{(\varrho)}(f) = D(D^{(\varrho-1)}(f))$, $\varrho \geqq 1$, falls $D^{(\varrho-1)}(f) \in \mathfrak{d}$. Ist $D^{(\varrho)}(f) \in \mathfrak{d}$, $\varrho \geqq 0$, so schreiben wir auch $f \in \mathfrak{d}^{(\varrho)}$; es soll also sein $\mathfrak{d} = \mathfrak{d}^{(0)}$ und aus $f \in \mathfrak{d}^{(r)}$, $r \geqq 0$, folgt, daß $D^{(r+1)}(f)$ existiert ($\in \mathfrak{d}'$) sowie daß $f \in \mathfrak{d}^{(\tau)}$, $0 \leqq \tau \leqq r$. — (2) Es sei $f_\nu(X) \in \mathfrak{d}^{(n-2)}$, $\nu = 1, \ldots, n$; $n \geqq 1$. Dann bezeichnen wir die Determinante [14] $|D^{(0)}(f_\nu) \ldots D^{(n-1)}(f_\nu)|$ mit $W_n(X)$ oder $W(f_1, \ldots, f_n)$ und als (die) Wronskische Determinante n-ter Ordnung aus den $f_1, \ldots, f_n$; allgemein bezeichnen wir für $1 \leqq k \leqq n$ als *Wronskische* (Unter-) Determinante k-ter Ordnung aus den $f_1, \ldots, f_n$ jede Determinante, die aus $W_n(X)$ durch Weglassen von beliebigen $(n-k)$-Zeilen $D^{(0)}(f_{\nu_\tau})$, $\ldots$, $D^{(n-1)}(f_{\nu_\tau})$, $\tau = 1, \ldots, n-k$, sowie der letzten $(n-k)$ Spalten $D^{(\mu)}(f_1), \ldots, D^{(\mu)}(f_n)$, $\mu = n-k, \ldots, n-1$, hervorgeht. — (3) Die Funktionen $f_\nu(X) \in \mathfrak{d}^{(n-2)}$, $\nu = 1, \ldots, n$; $n \geqq 1$, heißen linear abhängig im gemeinsamen Definitionsbereich $\mathfrak{M} \subset \mathfrak{R}$, wenn (von X „unabhängige" Konstante) $c_\nu \in \mathfrak{K}$ existieren, $\nu = 1, \ldots, n$, derart, daß die c_ν nicht sämtlich Null sind und daß $\sum c_\nu f_\nu(X) \equiv 0$ in $\mathfrak{M}$; solche c_ν heißen Abhängigkeitskoeffizienten (für die f_ν in $\mathfrak{M}$).

3. Identisches Verschwinden der Wronskischen Determinante.

Vorbemerkung. (a) Aus der Definition von $D^{(\varrho)}$ und $\mathfrak{d}^{(\varrho)}$, $\varrho \geqq 0$, folgt wegen Nr. 2. 1., (A_3) (I) vermöge vollständiger Induktion: $D^{(r)}(f_1 + \cdots + f_n) = D^{(r)}(f_1) + \cdots + D^{(r)}(f_n)$, falls $f_\nu \in \mathfrak{d}^{(r-1)}$. (b) Gemäß Nr. 2. 1., (A_4) (III) und (A_3) (I) sowie (II′) ist ferner $D^{(r)}(c_1 f_1 + \cdots + c_n f_n) = c_1 D^{(r)}(f_1) + \cdots + c_n D^{(r)}(f^n)$, wenn $c_\nu \in \mathfrak{K}$, $\nu = 1, \ldots, n$, und wenn der gemeinsame Definitionsbereich der f_ν sowohl *offen* als *zusammenhängend* ist.

Da in $\mathfrak{K}$ die übliche Determinantentheorie gilt [14], folgt (ohne Benutzung von Nr. 2. 1., (A_3) (II″)) der

1. Satz. Vor.[15] *Es sei* $\mathfrak{M}$ *offen und zusammenhängend* ($\mathfrak{M} \subset \mathfrak{R}$). *Ferner seien die* $f_\nu(X)$ $\mathfrak{M} \in \mathfrak{a}$, $\nu = 1, \ldots, n$; $n \geqq 1$, *linear abhängig in* $\mathfrak{M}$.

Beh. *Ist* $f_\nu(X) \in \mathfrak{d}^{(n-2)}$, $\nu = 1, \ldots, n$, *so gilt* $W(f_1, \ldots, f_n) \equiv 0$ *in* $\mathfrak{M}$.

In der Tat: Aus $f_\nu \in \mathfrak{d}^{(n-2)}$, $c_\nu \in \mathfrak{K}$, $\nu = 1, \ldots, n$, sowie aus $c_\nu \neq 0$ für mindestens ein ν und aus $F(X) = c_1 f_1(X) + \cdots + c_n f_n(X) \equiv 0$ folgt gemäß der Vorbemerkung (b): Es ist $F(X) \in \mathfrak{d}^{(n-2)}$ und $D^{(\varrho)}(F) \equiv 0 \equiv c_1 D^{(\varrho)}(f_1) + \cdots + c_n D^{(\varrho)}(f_n)$ in $\mathfrak{M}$, $\varrho = 0, \ldots, n-1$. Diese letzten

[14] Denn $\mathfrak{K}$ ist ein kommutativer Körper; vgl. z. B. Haupt, Einführung in die Algebra, Leipzig 1928, Nr. 9. 6., 9. 7.

[15] Der Satz gilt ohne die Voraussetzung, daß $\mathfrak{R}$ stark zusammenhängend im Kleinen ist in jedem seiner Punkte. Für $n = 1$ ist der Satz trivial. (Vgl. den Bew. von Satz 2.)

Gleichungen sind linear, homogen in den c_ν und besitzen nicht lediglich die triviale Lösung $c_1 = \cdots = c_n = 0$. Daher ist $W(f_1, \ldots, f_n) \equiv 0$ in $\mathfrak{M}$.

Unter gewissen Annahmen läßt sich umgekehrt aus dem Verschwinden der Wronskischen Determinante auf lineare Abhängigkeit schließen. Beispielsweise gilt der

2. Satz[15]. Vor. (1) *Es sei $\mathfrak{M}$ eine offene und zusammenhängende Teilmenge von $\mathfrak{R}$.* — (2) *Ferner sei $f_\nu(X) \mid \mathfrak{M} \in \mathfrak{d}^{(n-2)}$, $\nu = 1, \ldots, n$; $n \geqq 1$.* — (3) *Schließlich sei $W(f_1, \ldots, f_k) \equiv 0$ in $\mathfrak{M}$ für ein festes k mit $1 \leqq k \leqq n$.* — (4) *Im Falle $k > 1$ sei überdies $W(f_1, \ldots, f_{k-1}) \neq 0$ überall in $\mathfrak{M}$.*

Beh. *Die $f_1, \ldots, f_k$ (und erst recht also die $f_1, \ldots, f_n$) sind linear abhängig in $\mathfrak{M}$. Im Falle $k > 1$ ist das Verhältnis der Abhängigkeitskoeffizienten $c_1, \ldots, c_k$ eindeutig bestimmt, nämlich gleich dem Verhältnis der Wronskischen Unterdeterminanten $(k-1)$-ter Ordnung von $W(f_1, \ldots, f_k)$.*

Bew.[16]) Im Falle $k = 1$ ist $f_1(X) \equiv 0$ in $\mathfrak{M}$ (nach Vor. (3)), also die Beh. richtig. Es sei daher fortan $k \geqq 2$. — (**a**) Zunächst folgt, daß das Gleichungssystem

$$\text{(I)} \qquad \sum_{\varkappa=1}^{k} C_\varkappa(X)\, D^{(\varrho)}(f_\varkappa) \equiv 0, \quad 0 \leqq \varrho \leqq k-2,$$

in $\mathfrak{M}$ eine nicht-triviale Lösung besitzt. In der Tat: Eine solche Lösung ist z. B. $C_\varkappa(X) = W(f_1, \ldots, f_{\varkappa-1}, f_{\varkappa+1}, \ldots, f_k)$, $\varkappa = 1, \ldots, k$; also speziell $C_1(X) = W(f_2, \ldots, f_k)$, $C_k(X) = W(f_1, \ldots, f_{k-1})$. Da diese Wronskischen sämtlich von der Ordnung $(k-1)$ sind, da ferner $f_\varkappa \in \mathfrak{d}^{(k-2)}$ (gemäß Vor. (2)), so folgt aus Nr. 2. 1., (A_2) (III), daß $C_\varkappa(X) \in \mathfrak{d}$, $\varkappa = 1, \ldots, k$. Nach Vor. (4) ist überdies $C_k(X) \neq 0$ in $\mathfrak{M}$. — (**b**) Wegen $C_k(X) \neq 0$ in $\mathfrak{M}$ können die Gleichungen (I) durch $C_k(X)$ dividiert werden. Wir erhalten dadurch eine Lösung $C'_\varkappa(X) = C_\varkappa(X) : C_k(X)$, $\varkappa = 1, \ldots, k$, von (I) in $\mathfrak{M}$, wobei $C'_k(X) \equiv 1$ in $\mathfrak{M}$. Weil $\mathfrak{d}$ rational geschlossen ist (Nr. 2. 1., (A_3) (III)), gehören auch die $C'_\varkappa(X)$, $\varkappa = 1, \ldots, k-1$, zu $\mathfrak{d}$; und wegen Vor. (1) sowie Nr. 2. 1., (A_3) (II′) ist auch $C'_k(X) \in \mathfrak{d}$. — (**c**) Jetzt benutzen wir Vor. (3). Ihr zufolge ist $C'_1(X), \ldots, C'_k(X)$ Lösung auch von

$$\text{(II)} \qquad \sum_{\varkappa=1}^{k} C'_\varkappa(X)\, D^{(k-1)}(f_\varkappa) \equiv 0.$$

Im Hinblick auf die Vorbemerkung (b) ergibt sich aus (I) und (II), daß

$$\sum_{\varkappa=1}^{k} D(C'_\varkappa)\, D^{(\varrho)}(f_\varkappa) + \sum_{\varkappa=1}^{k} C'_\varkappa\, D^{(\varrho+1)}(f_\varkappa) \equiv \sum_{\varkappa=1}^{k-1} D(C'_\varkappa)\, D^{(\varrho)}(f_\varkappa) \equiv 0$$

für $0 \leqq \varrho \leqq k-2$.

[16]) Für den Fall reeller Funktionen einer reellen Veränderlichen vgl. auch Frechet, a. a. O.[2]).

Die Determinante des von den Gleichungen rechterhand gebildeten linearen, homogenen Gleichungssystems ist $W(f_1, \ldots, f_{k-1})$. Da diese Determinante in $\mathfrak{M}$ von Null verschieden ist (Vor. (4)) hat dieses Gleichungssystem in $\mathfrak{M}$ nur die triviale Lösung $D(C_\varkappa') \equiv 0, \varkappa = 1, \ldots, k-1$. Gemäß Vor. (1) und Nr. 2. 1., (A_3) (II″) ist somit $C_\varkappa'(X) \equiv$ konst $= c_\varkappa \in \mathfrak{K}$, $\varkappa = 1, \ldots, k-1$. Da überdies $C_k'(X) \equiv 1$ ist, ergibt sich die lineare Abhängigkeit der $f_1, \ldots, f_k$ in $\mathfrak{M}$ sowie die Eindeutigkeit von $v = c_1 : \cdots : c_k$.

Nunmehr läßt sich zeigen

3. Satz. Vor. *Es sei* $f_\nu(X) \mid \mathfrak{R} \in \mathfrak{d}^{(n-2)}, \nu = 1, \ldots, n$; $n \geqq 1$. *Ferner sei* $W(f_1, \ldots, f_n) \equiv 0$ *in* $\mathfrak{R}$.

Beh. (a) *Es existiert in* $\mathfrak{R}$ *ein System von zusammenhängenden, paarweise fremden, in* $\mathfrak{R}$ *offenen Mengen* $\mathfrak{Z}$, *deren Vereinigung* $\mathfrak{V}$ *(offen und) dicht in* $\mathfrak{R}$ *ist und wobei in jedem* $\mathfrak{Z}$ *die* $f_1, \ldots, f_n$ *linear abhängig sind.* — (b) *Eine Vereinigung* $\mathfrak{V}$ *der in* (a) *gekennzeichneten Art ist* $\mathfrak{V} = \mathfrak{R} - \mathfrak{Q}$, *wenn unter* $\mathfrak{Q}$ *die Vereinigung der Begrenzungen der Nullstellenmengen aller* Wronskischen *Determinanten (der Ordnungen 1 bis n) aus den* $f_1, \ldots, f_n$ *verstanden wird. Die Komponenten von* $\mathfrak{V}$ *sind Mengen* $\mathfrak{Z}$ *der in* (a) *genannten Art.*

Bew.[17]) (**1**) Es seien $W_{\nu\varrho}(X)$, $\varrho = 1, \ldots, \binom{n}{\nu}$, die sämtlichen Wronskischen Determinanten ν-ter Ordnung aus den $f_1, \ldots, f_n$ (in irgend einer Reihenfolge), $\nu = 1, \ldots, n$. Wegen $f_\nu \in \mathfrak{d}^{(n-2)}$ und Nr. 2. 1., (A_2) (II), sind die $f_\nu^{(\varkappa)}(X)$ stetig in $\mathfrak{R}$ für $0 \leqq \varkappa \leqq n-2$ und jedes ν; weil $\mathfrak{K}$ topologischer Körper ist (Nr. 2. 1., (A_1)), sind daher stetig in $\mathfrak{R}$ auch sämtliche $W_{\nu\varrho}(X)$ für $1 \leqq \nu \leqq n-1$. Es sei $\mathfrak{N}_{\nu\varrho}$ die Menge der Nullstellen von $W_{\nu\varrho}(X)$ in $\mathfrak{R}$; speziell ist $\mathfrak{N}_{n\varrho} = \mathfrak{N}_{n1} = \mathfrak{R}$, weil $W_n(X) \equiv 0$ in $\mathfrak{R}$. Zufolge der Stetigkeit der $W_{\nu\varrho}(X)$, $1 \leqq \nu \leqq n-1$, ist $\mathfrak{N}_{\nu\varrho}$ abgeschlossen in $\mathfrak{R}$ (für $\nu \leqq n-1$ und, wegen $\mathfrak{N}_{n1} = \mathfrak{R}$, auch für $\nu = n$). Daher ist der Rand $\mathfrak{Q}_{\nu\varrho} = \mathfrak{N}_{\nu\varrho} - \underline{\mathfrak{N}}_{\nu\varrho}$ (abgeschlossen und) nirgends dicht in $\mathfrak{R}$; mithin ist $\mathfrak{Q} = \bigcup_{\nu=1}^{n} \left(\bigcup_{\varrho=1}^{\binom{n}{\nu}} \mathfrak{Q}_{\nu\varrho} \right)$ ebenfalls abgeschlossen und nirgends dicht (in $\mathfrak{R}$). Folglich ist $\mathfrak{V} = \mathfrak{R} - \mathfrak{Q}$ offen und dicht in $\mathfrak{R}$, insbesondere ist also $\mathfrak{V} = \underline{\mathfrak{V}} \neq \mathfrak{O}$. — (**1**$_1$) Wir setzen noch $\mathfrak{S}_{\nu\varrho} = \underline{\mathfrak{N}}_{\nu\varrho}$ = offener Kern von $\mathfrak{N}_{\nu\varrho}$. Ist nun $\mathfrak{Z}'$ eine (nicht leere) zusammenhängende, sonst beliebige Teilmenge von $\mathfrak{V}$, so gilt für jedes $\mathfrak{S}_{\nu\varrho}$ *entweder* $\mathfrak{Z}'\mathfrak{S}_{\nu\varrho} = \mathfrak{O}$ *oder* $\mathfrak{Z}'(\mathfrak{R} - \mathfrak{S}_{\nu\varrho}) = \mathfrak{O}$ d. h. $\mathfrak{Z}' \subset \mathfrak{S}_{\nu\varrho}$. In der Tat: Wegen $\mathfrak{Z}' \subset \mathfrak{V}$ ist $\mathfrak{Z}'\mathfrak{Q}_{\nu\varrho} = \mathfrak{O}$ für jedes $\mathfrak{Q}_{\nu\varrho}$, also $\mathfrak{Z}'(\mathfrak{R} - \mathfrak{S}_{\nu\varrho}) = \mathfrak{Z}'(\mathfrak{R} - \mathfrak{N}_{\nu\varrho})$. Wäre nun $\mathfrak{Z}'(\mathfrak{R} - \mathfrak{N}_{\nu\varrho}) \neq \mathfrak{O}$ *und* $\mathfrak{Z}'\mathfrak{S}_{\nu\varrho} \neq \mathfrak{O}$, so existierte wegen $\mathfrak{Z}' = \mathfrak{Z}'\mathfrak{S}_{\nu\varrho} + \mathfrak{Z}'(\mathfrak{R} - \mathfrak{S}_{\nu\varrho}) = \mathfrak{Z}'\mathfrak{S}_{\nu\varrho} + \mathfrak{Z}'(\mathfrak{R} - \mathfrak{N}_{\nu\varrho})$ eine Spaltung von $\mathfrak{Z}'$ in zwei *fremde*, nicht leere, in $\mathfrak{Z}'$ *offene* Mengen, im Widerspruch damit, daß $\mathfrak{Z}'$ zusammenhängend sein soll. — (**1**$_2$) Ist

[17]) In Ziff. (1) dieses Beweises wird nicht benötigt, daß $\mathfrak{R}$ stark zusammenhängend im Kleinen ist. Für $n = 1$ ist der Satz trivial.

$\mathfrak{Z}' < \mathfrak{B}$, ist ferner $\mathfrak{Z}'$ nicht leer und zusammenhängend, sonst beliebig, so existiert zu $\mathfrak{Z}'$ (mindestens) ein m mit $1 \leqq m \leqq n$ und zu diesem m (mindestens) ein ϱ, etwa $\varrho = r$, mit $1 \leqq r \leqq \binom{n}{m}$ derart, daß $W_{mr}(X) \equiv 0$ in $\mathfrak{Z}'$ und daß, falls $m \geqq 2$, gleichzeitig alle $W_{m-1,\varrho}(X) \not\equiv 0$ sind für jedes $X \in \mathfrak{Z}'$ und jedes $\varrho = 1, \ldots, \binom{n}{m-1}$. In der Tat: Zu $\mathfrak{Z}'$ existiert ein kleinstes ν, etwa $\nu = m \geqq 1$, derart, daß $\mathfrak{Z}' < \mathfrak{S}_{m\varrho}$ für mindestens ein ϱ, etwa für $\varrho = r$; denn wegen $\mathfrak{Z}' < \mathfrak{S}_{n1} = \mathfrak{R}$ gibt es ein solches $m \leqq n$. Ist $m = 1$, so ist etwa $f_1(X) \equiv 0$ in $\mathfrak{Z}'$. Ist aber $m \geqq 2$, so folgt aus der Minimaleigenschaft von m und (1_1), daß $\mathfrak{Z}' \mathfrak{S}_{m-1,\varrho} = \mathfrak{D}$ für jedes $\varrho = 1, \ldots, \binom{n}{m-1}$. Wegen $\mathfrak{Z}' < \mathfrak{B}$ ist $\mathfrak{Z}' \mathfrak{D} = \mathfrak{D}$ und daher $\mathfrak{Z}' \mathfrak{R}_{m-1,\varrho} = \mathfrak{D}$ für jedes ϱ, also $W_{m-1,\varrho}(X) \not\equiv 0$ für jedes $X \in \mathfrak{Z}'$ und jedes ϱ, hingegen $W_{mr}(X) \equiv 0$ für jedes $X \in \mathfrak{Z}' < \mathfrak{S}_{mr}$.

(2) Ist $\mathfrak{Z}$ eine Komponente von $\mathfrak{B}$, so ist $\mathfrak{Z}$ selbst offen, weil $\mathfrak{R}$ *stark zusammenhängend im Kleinen* ist (gemäß Nr. 2. 2., (2)). Deswegen und wegen Ziff. (1_2) dieses Beweises ist Satz 2. auf $\mathfrak{Z}$ anwendbar; es sind also die $f_1, \ldots, f_n$ linear abhängig in $\mathfrak{Z}$. Da $\mathfrak{B}$ gleich der Vereinigung aller ihrer Komponenten ist, folgt die Beh. (a) und (b).

4. Hinreichende Bedingung für lineare Abhängigkeit.

Dem in Nr. 3., Satz 2., enthaltenen Abhängigkeitskriterium [15]) läßt sich das folgende an die Seite stellen (bei diesem ist (im Unterschied zu Nr. 3., Satz 2.,) die Voraussetzung wesentlich, daß $\mathfrak{R}$ stark zusammenhängend im Kleinen ist).

Satz. Vor. *Es sei $\mathfrak{M}$ eine offene und zusammenhängende Teilmenge von $\mathfrak{R}$. Ferner sei $f_\varkappa(X)$ $\mathfrak{M} \in \mathfrak{d}^{(k-2)}$, $\varkappa = 1, \ldots, k$; $1 \leqq k$, sowie $W(f_1, \ldots, f_k) \equiv 0$ in $\mathfrak{M}$. Falls $k \geqq 2$ ist, sei schließlich in jedem Punkt $X \in \mathfrak{M}$ je mindestens eine der Wronskischen Determinanten $(k-1)$-ter Ordnung aus den $f_1, \ldots, f_k$ von Null verschieden.*

Beh. (**a**) *Die $f_1, \ldots, f_k$ sind linear abhängig in $\mathfrak{M}$.* — (**b**) *Für jede der Wronskischen Determinanten $(k-1)$-ter Ordnung $W(f_{r_1}, \ldots, f_{r_{k-1}})$ aus den $f_1, \ldots, f_k$ gilt: Entweder ist $W(f_{r_1}, \ldots, f_{r_{k-1}}) = 0$ überall in $\mathfrak{M}$ oder es ist $W(f_{r_1}, \ldots, f_{r_{k-1}}) \not\equiv 0$ überall in $\mathfrak{M}$, wobei $k \geqq 2$.*

Bew. (**1**) Es sei O. B. d. A. $k \geqq 2$, ferner sei $W_{k-1}(X) = W(f_1, \ldots, f_{k-1})$ gesetzt und es sei z. B. $W_{k-1}(Y) \not\equiv 0$ für ein $Y \in \mathfrak{M}$. Wegen Nr. 2. 1., (A_2) (II), ist $W_{k-1}(X)$ stetig in Y (vgl. auch Nr. 3., Satz 3., Bew., Ziff. (1)), also $W_{k-1}(X) \not\equiv 0$ in einer Umgebung $\mathfrak{U}$ von Y. Ist $\mathfrak{T} = \mathfrak{T}(Y)$ die Y enthaltende Komponente von $\mathfrak{U}\mathfrak{M}$, so ist $\mathfrak{T}$ (nicht leer und) offen (gemäß Nr. 2. 2., (2), Satz); es sind also (gemäß Nr. 3., Satz 2.) die $f_1, \ldots, f_k$ linear abhängig in $\mathfrak{T}$ und zwar *mit eindeutig bestimmtem Verhältnis v der Abhängigkeitskoeffizienten*, falls $k \geqq 2$.

(2) Jeder Punkt $X \in \mathfrak{M}$ ist in einer *größten* (offenen oder) abgeschlossenen, zusammenhängenden (nicht leeren) Teilmenge $\mathfrak{Z}$ von $\mathfrak{M}$ enthalten, in der die $f_1, \ldots, f_k$ *linear abhängig* sind mit *eindeutig bestimmtem* v (falls $k \geqq 2$). In der Tat: *Erstens* existieren offene zusammenhängende (nicht leere), X enthaltende Teilmengen $\mathfrak{Z}'$ von $\mathfrak{M}$, in denen die $f_1, \ldots, f_k$ linear abhängig sind; z. B. ist die in Ziff. (1) eingeführte Menge $\mathfrak{T}(X)$ ein solches $\mathfrak{Z}'$. *Zweitens* ist die Vereinigung $\mathfrak{Z}^*$ beliebig vieler derartiger Mengen $\mathfrak{Z}'$ (mit $X \in \mathfrak{Z}'$), in deren jeder die $f_1, \ldots, f_k$ linear abhängig sind, wieder ein derartiges $\mathfrak{Z}'$; denn $\mathfrak{Z}^*$ ist offen und zusammenhängend, ferner enthält (wegen $X \in \mathfrak{Z}'$ und gemäß Ziff. (1) dieses Beweises) jedes $\mathfrak{Z}'$ eine zusammenhängende Umgebung von X, in der v eindeutig bestimmt ist falls $k \geqq 2$; deswegen und weil im Durchschnitt zweier solcher Umgebungen wieder eine derartige Umgebung enthalten ist, besitzt v für alle $\mathfrak{Z}'$ den gleichen Wert. *Drittens* sind wegen der Stetigkeit der $f_\varkappa \in \mathfrak{d}$ (gemäß Nr. 2. 1., (A_2) (II)) die $f_1, \ldots, f_k$ linear abhängig, wenn in $\mathfrak{Z}^*$, dann auch in der abgeschlossenen Hülle $\overline{\mathfrak{Z}}^*$ von $\mathfrak{Z}^*$ und zwar mit dem gleichen v wie in $\mathfrak{Z}^*$.

(3) Es sei nun $\mathfrak{Z}$ abgeschlossene Hülle einer größten offenen, zusammenhängenden Menge $\mathfrak{Z}^* < \mathfrak{M}$ (im Sinne von Ziff. (2) dieses Beweises). Zum Beweise der Beh. (a) genügt es zu zeigen, daß $\mathfrak{Z} = \mathfrak{M}$ ist; da alsdann (für $k \geqq 2$) das Verhältnis der Abhängigkeitskoeffizienten einerseits in $\mathfrak{M}$ eindeutig bestimmt ist und andererseits (gemäß Nr. 3., Satz 2.) in jedem Punkt von $\mathfrak{M}$ gleich dem der Wronskischen Unterdeterminanten $(k-1)$-ter Ordnung von $W(f_1, \ldots, f_k)$ ist, folgt dann auch die Beh. (b). Angenommen nun, es sei $\mathfrak{M}' = \mathfrak{M} - \mathfrak{Z} \neq \mathfrak{O}$. Da $\mathfrak{M}$ zusammenhängend sein soll, müßte dann $\overline{\mathfrak{M}}' \mathfrak{Z} \neq \mathfrak{O}$ sein. Es würde somit ein Punkt $Y \in \mathfrak{Z}$ existieren derart, daß Y Häufungspunkt von Punkten $H \in \mathfrak{M}'$ wäre. Es sei (vgl. Vor.) etwa $W_{k-1}(Y) \neq 0$ und $\mathfrak{U}$ eine zusammenhängende Umgebung von Y mit $W_{k-1}(X) \neq 0$ für $X \in \mathfrak{U}$ (vgl. Ziff. (1) dieses Beweises sowie Nr. 2. 1., (A_1)). Es existierte nun ein $H \in \mathfrak{M}' \mathfrak{U}$ und es wäre $\mathfrak{U}$ eine offene zusammenhängende (nicht leere) Menge mit $Y \in \mathfrak{U}$ und $H \in \mathfrak{U}$. Wegen $H \in \mathfrak{M}'$ und $\mathfrak{M}' \mathfrak{Z} = \mathfrak{O}$ wäre $\mathfrak{Z}'' = \mathfrak{Z} \cup \mathfrak{U} \neq \mathfrak{Z}$. Andererseits wären die $f_1, \ldots, f_k$ linear abhängig in $\mathfrak{Z}''$ und zwar (falls $k \geqq 2$) mit dem gleichen v wie in $\mathfrak{Z}$ (gemäß Ziff. (2) dieses Beweises). Wegen $\mathfrak{Z}' < \mathfrak{Z}''$, $\mathfrak{Z}' \neq \mathfrak{Z}''$ wäre somit $\mathfrak{Z}$ nicht maximal entgegen der Definition von $\mathfrak{Z}$. Wir haben also $\mathfrak{Z} = \mathfrak{M}$, w. z. z. w.

5. Abänderung des Axiomensystems in Nr. 2. 1.

Die Sätze in Nr. 3. und Nr. 4. bleiben mutatis mutandis richtig, wenn man einerseits von $\mathfrak{R}$ etwas weniger fordert, andererseits das Mengensystem $\mathfrak{t}$ vergrößert. Im einzelnen handelt es sich um folgende

Modifikationen: In Nr. 2. 1., (A_1) werde $\mathfrak{R}$ lediglich als topologischer Raum (ohne Trennungsaxiome) postuliert, hingegen $\mathfrak{t}$ als das System *aller* Teilmengen von $\mathfrak{R}$ angenommen. [In (A_2) (III) ist dann $\mathfrak{T}'_0$ nicht mehr notwendig offen in $\mathfrak{R}$]. In (A_3), (II′) und (II″) wird $\mathfrak{M}$ bzw. $\mathfrak{Q}$ lediglich als zusammenhängend (nicht als offen in $\mathfrak{R}$) angenommen.

Demgemäß entfällt Nr. 2. 2., (2). Ferner ist in Nr. 3., Satz 1. und 2. die Menge $\mathfrak{M}$ lediglich als zusammenhängend vorauszusetzen. In Satz 3., Beh. (a) kann von den $\mathfrak{Z}$ nicht mehr (ohne weiteres) gesagt werden, daß sie offen sind in $\mathfrak{R}$. In Nr. 4., Satz., Vor., ist $\mathfrak{M}$ (nicht als offen aber als) zusammenhängend und zugleich als *zusammenhängend im Kleinen* anzunehmen. (Der Beweis dieses Satzes bleibt richtig, wenn die dort auftretenden Mengen $\mathfrak{T}(Y)$, $\mathfrak{Z}'$ usw. nur mehr zusammenhängend, nicht zugleich offen sind.)

(Eingegangen am 6. November 1949.)

Mathematische Zeitschrift, Band 53, Heft 2, S. 131—132 (1950).

Über die Gaußsche Krümmung der Überraumflächen.

Herrn Erich Kamke zum 60. Geburtstag am 18. August 1950 gewidmet.

Von

E. Bompiani.

1. Es ist bekannt, daß man für jeden Punkt einer einem euklidischen Raum S_n angehörenden Fläche eine metrische Invariante, die Gausssche Krümmung, definieren kann. Während jedoch für $n = 3$ diese eine sofort klare geometrische Bedeutung mittels des Begriffs der Krümmungsradien der Fläche im betrachteten Punkt hat, mangelt diese Deutung für $n > 3$. Eine einfache Auslegung für ein beliebiges n ist der Inhalt dieser Abhandlung.

2. Wenn die Fläche in parametrischer Darstellung mit orthogonalen monometrischen Koordinaten durch

$$x^i = x^i(u, v) \qquad i = 1, 2, \ldots, n \tag{2.1}$$

gegeben ist, wird ihre Krümmung mittels der Discriminanten folgender zwei Formen ausgedrückt: die erste

$$ds^2 = E\,du + 2F\,du\,dv + G\,dv^2 \tag{2.2}$$

mit den üblichen Ausdrücken der Koeffizienten ($E = \sum_1^n (x_u^i)^2, \ldots$) und die zweite, für $n > 3$, rein *symbolischer* Natur:

$$L\,du^2 + 2M\,du\,dv + N\,dv^2 \tag{2.3}$$

deren Koeffizienten die Matrizen [1])

$$L = (x_{uu}, x_u, x_v);\ M = (x_{uv}, x_u, x_v);\ N = (x_{vv}, x_u, x_v) \tag{2.4}$$

sind, die mit den Koordinaten x_u^i, x_v^i usw. gebildet werden.

Gesagte Krümmung ist

$$K = \frac{LN - M^2}{(EG - F^2)^2} \tag{2.5}$$

[1]) Der symbolischen Differenzialformen (die als Koeffizienten Matrizen haben) habe ich mich in metrischen Problemen bezüglich der Überflächen schon in anderen Arbeiten bedient. Man siehe z. B. die vier Abhandlungen: *Invarianti e covarianti metrici nelle deformazioni di specie superiore delle superficie* (Rend. Acc. Lincei, s. V, vol. XXIX e XXX, 1920—21) und *Geometria riemanniana di specie superiore* (Memorie Acc. d'Italia, vol. VI, n. 8, 1935).

deren Invarianz für Koordinatenänderung (gewollter Art) und Parameterdarstellung sofort bewiesen werden kann.

Wenn $n > 3$, können die Produkte der Matrizen des Nenners als Summe der Produkte der Determinanten 3. Ordnung erhalten werden, die in ihnen enthalten sind und dem gleichen Indextripel i, h, l, (voneinander verschieden) entsprechen.

Wenn man nur die Koordinaten x^i, x^h, x^l betrachtet, so bedeutet dies, die orthogonale Projektion der Fläche auf den S_3 der Axen x^i, x^h, x^k ins Auge zu fassen. Die Krümmung dieser Projektion in der Projektion eines fixierten Punktes bezeichne man mit K_{ihl}; sie ist durch die auf diesen Fall angewandte Formel 2.5 durch

$$2.6. \qquad K_{ihl} \cdot \Delta^2_{ihl} = \begin{vmatrix} x^i_{uu} & x^h_{uu} & x^l_{uu} \\ x^i_u & x^h_u & x^l_u \\ x^i_v & x^h_v & x^l_v \end{vmatrix} \begin{vmatrix} x^i_{vv} & x^h_{vv} & x^l_{vv} \\ x^i_u & x^h_u & x^l_u \\ x^i_v & x^h_v & x^l_v \end{vmatrix} - \begin{vmatrix} x^i_{uv} & x^h_{uv} & x^l_{uv} \\ x^i_u & x^h_u & x^l_u \\ x^i_v & x^h_v & x^l_v \end{vmatrix}^2$$

gegeben, wo

$$2.7 \qquad \Delta^2_{ihl} = \begin{pmatrix} x^i_u & x^h_u & x^l_u \\ x^i_v & x^h_v & x^l_v \end{pmatrix}^2$$

bedeutet.

Dieser letzte Ausdruck, mit $du^2 dv^2$ multipliziert, gibt das Quadrat des Flächenelements relativ der betrachteten Projektion; wogegen $\sqrt{EG - F^2}\, du dv$ das Flächenelement der gegebenen Fläche ist.

Man hat somit

$$K = \sum_{ihl} K_{ihl} \frac{\Delta^2_{ihl}}{(EG - F^2)^2}$$

was man in folgender Art liest:

Nachdem man willkürlich in einem Punkt einer Fläche des S_n, n durch ihn laufende, aufeinander senkrechte Geraden gewählt hat, projiziere man die Fläche orthogonal auf jeden der durch drei jener Geraden bestimmten S_3. Die Summe der mit der vierten Potenz des Verhältnisses der auf den Projektionen stehenden Flächenelemente zu denen der gegebenen Fläche selbst multiplizierten Krümmungen der Projektionen ist die Gaussssche *Krümmung der Fläche.*

Die Anzahl jener Projektionen kann um zwei vermindert werden, indem man zwei jener n Geraden in der Tangentialebene nimmt; tatsächlich sind die einzigen $\Delta_{ihl} \neq 0$ die $\Delta_{12l} = 1$. Daher:

Man nehme $n - 2$ Geraden durch den Punkt, senkrecht aufeinander und auf die Fläche und man projiziere die letztere orthogonal auf die die Tangentialebene und die einzelnen genannten Normalen enthaltende S_3: die Summe der Gaussschen *Krümmungen der Projektionen ist die* Gaussssche *Krümmung der Fläche.*

(Eingegangen am 18. November 1949.)

Mathematische Zeitschrift, Band 53, Heft 2, S. 133—135 (1950).

Eine Normalform für endliche rein-inseparable Körpererweiterungen.

Herrn Erich Kamke zum 60. Geburtstag am 18. August 1950 gewidmet.

Von

Günter Pickert in Tübingen.

In einer früheren Arbeit[1]) habe ich eine Normalform für endliche rein-inseparable Körpererweiterungen aufgestellt, doch gelang mir der Beweis für die Invarianz gewisser dabei auftretender Exponenten nur auf dem Umweg über das Radikal einer primären kommutativen Algebra. Auch in einer weiteren Arbeit[2]) konnte ich nur den Beweis für die Existenz der Normalform etwas vereinfachen, jedoch keine Abkürzung des Invarianzbeweises geben. Kürzlich hat mich nun im Zusammenhang mit der genannten Normalform Herr J. Dieudonné auf eine von ihm benutzte Methode[3]) hingewiesen. Mit dieser gelingt tatsächlich ein sehr einfacher Beweis sowohl für die Existenz der Normalform wie für die Invarianz der darin auftretenden Exponenten. Diesen Beweis stelle ich im folgenden dar.

Es sei K ein Körper der Charakteristik $p > 0$. Dann lautet der zu beweisende Satz:

a) *Je m Erzeugende einer rein-inseparablen Erweiterung $L \mid K$ der Vielfachheit m lassen sich so als $a_1, \ldots, a_m$ bezeichnen, daß mit $i = 1, \ldots, m$ die Bedingungen*

I. $a_i^{q_i} = \alpha_i \in K(a_1^{q_i}, \ldots, a_{i-1}^{q_i}) \quad (q_i = p^{e_i}, e_i > 0)$

II. $a_i^{p^{e_i-1}}$ *nicht in* $K(a_1, \ldots, a_{i-1})$

III. $e_1 \geqq e_2 \geqq \cdots \geqq e_m$

erfüllt sind.

b) *Gelten umgekehrt für eine Erzeugung $L = K(a_1, \ldots, a_m)$ einer rein-inseparablen Erweiterung $L \mid K$ die Bedingungen* I *und* II, *so sind die Exponenten $e_1, \ldots, e_m$ bis auf die Reihenfolge Invarianten der*

[1]) G. Pickert: Neue Methoden i. d. Strukturtheorie d. komm. assoz. Algebren Math. Ann. 116 (1939), S. 217—280, insb. S. 220.

[2]) G. Pickert: Inseparable Körpererweiterungen, Math. Zeitschrift 52 (1949) S. 81—135, insb. S. 90, 91. Auf den § 1 dieser Arbeit sei hinsichtlich der im folgenden verwandten Begriffe verwiesen.

[3]) J. Dieudonné: Sur les extensions transcendantes séparables, Summa Brasiliensis Mathematicae II (1947), S. 1—20, insb. S. 14.

Erweiterung $L|K$ und m die Vielfachheit dieser Erweiterung. Die $a_1, \ldots, a_m$ lassen sich dann noch so umordnen, daß mit denselben e_i neben I *und* II *auch noch* III *gilt.*

Zum Beweis von Teil a) sei e der Exponent und m die Vielfachheit der rein-inseparablen Erweiterung $L|K$ sowie A eine Menge von m Elementen mit $L = K(A)$. Der Methode von DIEUDONNÉ[3]) folgend betrachte ich die Körper $L_k = K(L^{p^k}) = K(A^{p^k})$ $(k = 0, 1, \ldots, e)$, bei dennen also $L_0 = L$ und $L_e = K$ ist. Rekursiv werden die Mengen A_k $(k = 0, 1, \ldots, e)$ gebildet gemäß den Bedingungen: A_{k+1} ist für $k = 0, 1, \ldots, e-1$ eine minimale Untermenge von A_k mit

$$L_{k+1} = K(A_{k+1}^{p^{k+1}}),$$

während $A_0 = A$ gesetzt wird. Wegen $L_{k+1} = K(L_k^p)$ lassen sich solche Mengen A_k bestimmt angeben. Dabei ist natürlich A_e stets leer und es gilt $A_0 \supseteq A_1 \supseteq \cdots \supseteq A_{e-1}$. Ich numeriere nun die Elemente von A so, daß $A_k = \{a_1, \ldots, a_{m_k}\}$ ist. Dabei besteht zwischen den m_k die Beziehung

$$m = m_0 \geqq m_1 \geqq \cdots \geqq m_{e-1} > m_e = 0.$$

Als e_i wird jetzt diejenige natürliche Zahl $k \leqq e$ bezeichnet, für welche $a_i \in A_{k-1} - A_k$ gilt. Die $e_1, \ldots, e_m$ sind dann also alle positiv und erfüllen die Bedingung III. Mit $q_i = p^{e_i}$ ist weiter wegen

$$K(a_1^{q_i}, \ldots, a_{i-1}^{q_i}) \supseteq K(A_k^{p^k}) = L_k \qquad (\text{für } e_i = k)$$

auch Bedingung I erfüllt. Daraus, daß $A_k^{p^k}$, aber keine echte Untermenge davon die Erweiterung $L_k|K$ erzeugt, folgt in bekannter Weise[4])

$$p^{m_k} = (L_k : L_k^p(K)) = (L_k : L_{k+1}).$$

Daraus folgt

$$p^{\sum_{k=0}^{e-1} m_k} = \prod_{k=0}^{e-1} (L_k : L_{k+1}) = (L : K),$$

also

$$\sum_{k=0}^{e-1} m_k = f, \tag{1}$$

wenn $(L : K) = p^f$ gesetzt wird. Bezeichnet jetzt einmal e_i' den Exponenten von a_i bez. $K(a_1, \ldots, a_{i-1})$, so gilt wegen I die Ungleichung

$$e_i \geqq e_i'. \tag{2}$$

Andererseits bestehen die Gleichungen

$$\sum_{i=1}^{m} e_i' = f \tag{3}$$

und

$$\sum_{i=1}^{m} e_i = \sum_{k=1}^{e} k(m_{k-1} - m_k) = \sum_{k=0}^{e-1} m_k. \tag{4}$$

[4]) Siehe Fußnote [2]) a. a. O. S. 88.

Aus (1) bis (4) folgt nun $e_i' = e_i$, so daß auch Bedingung II erfüllt ist. Dies beendet den Beweis von Teil a).

Sei jetzt andererseits $L = K(a_1, \ldots, a_m)$ eine den Bedingungen I und II genügende Erzeugung und e wieder der Exponent von $L|K$. Ist dann B_k die Menge der a_i mit $e_i > k$ $(k = 0, 1, \ldots, e)$, so gilt

$$L_k = K(L^{p^k}) = K(B_k^{p^k});$$

denn gäbe es eine kleinste Zahl i mit $e_i \leqq k$ so, daß $a_i^{p^k}$ nicht in $K(B_k^{p^k})$ liegt, so würde aus I der Widerspruch

$$a_i^{p^k} \in K(a_1^{p^k}, \ldots, a_{i-1}^{p^k}) \leqq K(B_k^{p^k}) \text{ folgen.}$$

Mit n_k als Elementezahl von B_k und m_k als Vielfachheit von $L_k|K$ ergibt sich daher

$$m_k \leqq n_k. \tag{5}$$

Da e_i wegen I und II der Exponent von a_i bez. $K(a_1, \ldots, a_{i-1})$ ist, gilt

$$\sum_{i=1}^{m} e_i = f, \tag{6}$$

wobei wieder $(L:K) = p^f$ gesetzt sei. Weil $n_{k-1} - n_k$ $(k = 1, \ldots, e)$ die Anzahl der natürlichen Zahlen $i \leqq m$ mit $e_i = k$ ist und natürlich $e \geqq e_i$ $(i = 1, \ldots, m)$ gilt, erhält man ganz entsprechend zu (4) die Gleichung

$$\sum_{i=1}^{m} e_i = \sum_{k=0}^{e-1} n_k. \tag{7}$$

Aus (1), (5), (6) und (7) folgt nun $m_k = n_k$. Die $e_1, \ldots, e_m$ sind also durch die invariant definierten m_k bis auf ihre Reihenfolge eindeutig bestimmt. Insbesondere ist $m = m_0$, d. h. m die Vielfachheit von $L|K$. Wegen $m_k = n_k$ und $B_0 \geqq B_1 \geqq \cdots \geqq B_{e-1} > B_e = 0$ können nun die B_k als Mengen A_k bei der Aufstellung der Normalform gemäß dem Beweis von Teil a) genommen werden. Die dort vorgenommene Numerierung ergebe die Reihenfolge $a_{\varphi(1)}, \ldots, a_{\varphi(m)}$. Bei der so erhaltenen, I—III genügenden Normalform gehört nun nach Konstruktion zu $a_{\varphi(i)}$ als Exponent (bez. $K(a_{\varphi(1)}, \ldots, a_{\varphi(i-1)})$) die Zahl k mit $a_{\varphi(i)} \in B_{k-1} - B_k$, und nach Definition von B_k ist dann $k = e_{\varphi(i)}$. Damit sind alle Behauptungen von Teil b) bewiesen. Weil nach dem Beweis von Teil a) ohne Zerstörung der Normalformeigenschaften I—III die Elemente innerhalb jeder Menge $A_{k-1} - A_k$ noch beliebig umgeordnet werden dürfen, kann für φ jede Permutation der Ziffern $1, \ldots, m$ genommen werden, welche der Bedingung

$$e_{\varphi(1)} \geqq e_{\varphi(2)} \geqq \cdots \geqq e_{\varphi(m)}$$

genügt.

(Eingegangen am 30. Dezember 1949.)

Mathematische Zeitschrift, Band 53, Heft 2, S. 136—148 (1950).

Asymptotische Eigenschaften der unvollständigen Gammafunktion.

Herrn Erich Kamke zum 60. Geburtstag am 18. August 1950 gewidmet.

Von

F. G. Tricomi z. Z. in Pasadena, California.

1. Seit einiger Zeit pflege ich die unvollständige Gammafunktion

$$\gamma(\alpha, x) = \int_0^x e^{-t} t^{\alpha-1} dt \tag{1}$$

das *Aschenbrödel der Funktionen* zu nennen, weil es mir scheint, daß sie trotz ihres unbestreitbaren Interesses bis jetzt zu wenig beachtet worden ist. Selbst ich habe sie bis jetzt verkannt, bis die Redaktion eines Berichtes über diesen Gegenstand für das *Bateman Manuscript Project*[1]) mir gezeigt hat, wie wichtig diese Funktion ist, um u. a. eine einheitliche Betrachtung vieler Funktionen zu ermöglichen: Fehlerintegral, Integral-Sinus und -Kosinus, Exponential-Integral, Fresnelsche Integrale usw., die sonst recht vereinzelt in der Literatur erscheinen.

Im Laufe jener Arbeit habe ich eine Menge von vermutlich neuen Eigenschaften der unvollständigen Gammafunktion entdeckt, über die hier ausführlich zu berichten nicht möglich ist. Darum werde ich mich darauf beschränken, im wesentlichen hier nur über einige hübsche asymptotische Eigenschaften unserer Funktion zu sprechen, welche u. a. die bis jetzt kaum beachteten Nullstellen der Funktion $\gamma(\alpha, x)$ zu beherrschen erlauben.

2. Besser als auf die Integraldarstellung (1) — welche nur für $\Re\alpha > 0$ gültig ist — läßt sich die Theorie der unvollständigen Gammafunktion auf die konfluente hypergeometrische Funktion

$$\gamma^*(\alpha, x) = e^{-x}\, {}_1F_1(1, \alpha+1; x) = \frac{e^{-x}}{\Gamma(\alpha+1)} = e^{-x} \sum_{n=0}^{\infty} \frac{x^n}{\Gamma(\alpha+n+1)} \tag{2}$$

gründen, welche eine *ganze* Funktion ihrer *beiden* Argumente α und x ist. Dann lassen sich die Funktion $\gamma(\alpha, x)$, ihr *Komplement* $\Gamma(\alpha, x)$ und die ebenso oft nützliche Funktion

$$\gamma_1(\alpha, x) = e^{-\alpha\pi i}\gamma(\alpha, e^{\pi i}x) = e^{\alpha\pi i}\gamma(\alpha, e^{-\pi i}x) \tag{3}$$

[1]) Ein vom Office of Naval Research der USA. unterstütztes Unternehmen des California Institute of Technology in Pasadena.

leicht durch die Gleichungen

$$\gamma(\alpha, x) = \Gamma(\alpha)\, x^{\alpha}\, \gamma^*(\alpha, x), \quad \gamma_1(\alpha, x) = \boldsymbol{\Gamma}(\alpha)\, x^{\alpha}\, \gamma^*(\alpha, -x), \tag{4}$$
$$\boldsymbol{\Gamma}(\alpha, x) = \Gamma(\alpha) - \gamma(\alpha, x)$$

definieren, wo die Potenz x^{α} ihren Hauptwert

$$\text{H.W. } x^{\alpha} = |x|^{\alpha}\, e^{i\alpha \arg x} \qquad (-\pi < \arg x \leqq \pi)$$

haben soll.

Die Bestimmung des asymptotischen Verhaltens dieser Funktionen, wenn $|\alpha|$ ins Unendliche wächst, *bei beschränkt bleibendem* $|x|$, bietet keine Schwierigkeit dar, weil die Grundreihe (2) unter diesen Umständen selbst einen asymptotischen Charakter besitzt. Diese Frage hat aber auch nur ein geringes Interesse. Was dagegen am meisten interessant aber auch schwierig ist, ist zu untersuchen, was geschehen wird, wenn $|\alpha|$ und $|x|$ *gleichzeitig* wachsen: eine Frage welche im Falle der allgemeinen konfluenten hypergeometrischen Funktionen bis jetzt nur relativ wenig untersucht worden ist[2]), abgesehen von dem Sonderfalle der Laguerreschen Polynome[3]).

In dieser Untersuchung sind — wie es in solchen Fragen öfter vorkommt — zwei Fälle zu unterscheiden: Ein *allgemeiner Fall* bei dem α und x nicht annähernd gleich sind, und ein *spezieller*, schwierigerer Fall, wenn x in der Nähe von $\alpha - 1$ liegt. Wir wollen sie nacheinander betrachten.

3. Im *Allgemeinen Fall* gehen wir von der Integraldarstellung der Komplement-Funktion

$$\Gamma(\alpha + 1, x) = \int_x^{\infty} e^{-t}\, t^{\alpha}\, dt$$

aus, welche — im Gegensatz zu (1) — für jedes nicht reelle-negative x gültig ist, solange der Integrationsweg die negative reelle Achse vermeidet und so ins Unendliche geht, daß darauf $\lim \Re t = +\infty$ ist. Noch besser, wenn wir die Substitution $t = x(1 + \tau)$ ausüben, können wir

$$\Gamma(\alpha + 1, x) = e^{-x}\, x^{\alpha+1} \int_0^{\infty e^{i\psi}} e^{-x\tau}\, (1 + \tau)^{\alpha}\, d\tau$$

schreiben, wo die Integration entlang einem passenden Strahl: $\arg \tau = \psi$ auszuführen ist.

---

[2]) Literaturangaben darüber findet man in meiner Arbeit: Sulle funzioni ipergeometriche confluenti, Annali di Matem. (4) **26** (1947), 141—175. Dazu kommt noch: W. C. Taylor, A complete set of asymptotic formulas..., Journ. Math. Ph. M. I. T. **18** (1939), 34—49.

[3]) S. meine zusammenfassende Arbeit: Sul comportamento asintotico dei polinomi di Laguerre, Annali di Matem. (4) **28** (1949).

Um das letzte Integral asymptotisch abzuschätzen, spalten wir den Integrationsweg in die beiden Teile $(0, t_0)$ und $(t_0, \infty e^{i\psi})$ auf, wo

$$t_0 = \vartheta e^{i\psi} \qquad (0 < \vartheta < 1)$$

einen willkürlichen aber festen Punkt unseres Strahles innerhalb des Einheitskreises bedeutet. Außerdem führen wir durch Multiplizieren und Dividieren mit $e^{\alpha\tau}$ und nachträglicher Änderung von τ in t, unter dem Integralzeichen die elementare Funktion

$$F(t) = e^{-\alpha t}(1+t)^{\alpha}$$

ein. So erhalten wir

$$\Gamma(\alpha+1, x) = e^{-x} x^{\alpha+1}(H_1 + H_2) \tag{5}$$

wo

$$H_1 = \int_0^{t_0} e^{(\alpha-x)t} F(t)\, dt, \qquad H_2 = \int_{t_0}^{\infty e^{i\psi}} e^{(\alpha-x)t} F(t)\, dt.$$

Wir fangen mit der Untersuchung von H_1 an.

Da $|t_0| < 1$ ist, dürfen wir die Funktion $F(t)$ nach Potenzen von t entwickeln und so ergibt sich

$$H_1 = \sum_{m=0}^{\infty} \frac{F^{(m)}(0)}{m!} \int_0^{t_0} e^{(\alpha-x)t} t^m\, dt = \sum_{m=0}^{\infty} \frac{F^{(m)}(0)}{(x-\alpha)^{m+1}} \{1 - e^{(\alpha-x)t_0} e_m[(x-\alpha)t_0]\},$$

wo $e_m(z)$ die $(m+1)$-te Teilsumme der Exponentialreihe:

$$e_m(z) = \sum_{h=0}^{m} \frac{z^h}{h!},$$

d. h. ein gewisses Polynom m-ten Grades bedeutet. Aber, solange $\Re[(\alpha - x)\, t_0] < 0$, d. h. solange

$$\cos(\xi + \psi) > 0, \tag{6}$$

wo ξ das Argument von $x - \alpha$ bedeutet, streben die Glieder mit dem Faktor $e^{(\alpha-x)t_0}$ *exponentiell* gegen Null wenn $|x-\alpha|$ ins Unendliche rückt; unter diesen Umständen haben wir also

$$H_1 \sim \sum_{m=0}^{\infty} \frac{F^{(m)}(0)}{(x-\alpha)^{m+1}} \tag{7}$$

(im Sinne von Poincaré), wenigstens solange α (wovon $F^{(m)}(0)$ abhängt) bei wachsendem $|x-\alpha|$ festbleibt.

Allgemeiner können wir behaupten, daß die asymptotische Entwicklung (7) sogar für ein wachsendes $|\alpha|$ gilt, solange die Reihe (7) ihren asymptotischen Charakter bewahrt, d. h. solange

$$\lim \frac{F^{(m)}(0)}{(x-\alpha)^{m+1}} : \frac{F^{(m-1)}(0)}{(x-\alpha)^{m}} = 0$$

ist, was sicher der Fall, ist wenn

$$|\alpha| = o(|x - a|^2) \tag{8}$$

ist, weil — wie wir jetzt sehen werden — $F^{(m)}(0)$ *ein Polynom* $[m/2]$*-ten Grades in* α *ist.*

In der Tat[4]) haben wir

$$F(t) = \exp\{\alpha[\log(1+t) - t]\} = \exp\left\{-\alpha t^2\left(\frac{1}{2} - \frac{t}{3} + \frac{t^2}{4} - \cdots\right)\right\}$$
$$= \sum_{n=0}^{\infty} \frac{(-\alpha)^n}{n!} t^{2n}\left(\frac{1}{2} - \frac{t}{3} + \frac{t^2}{4} - \cdots\right)^n$$

und infolgedessen, wenn wir allgemein

$$\left(\frac{1}{2} + \frac{z}{3} + \frac{z^2}{4} + \cdots\right)^n = \sum_{k=0}^{\infty} c_k^{(n)} z^k \qquad (n = 0, 1, 2, \ldots)$$

setzen[5]), wobei

$$c_0^{(n)} = \frac{1}{2^n}, \qquad c_1^{(n)} = \frac{n}{2^{n-1} \cdot 3}, \qquad c_2^{(n)} = \frac{n(4n+5)}{2^{n+1} \cdot 3^2}, \ldots,$$

können wir schreiben

$$F(t) = \sum_{n=0}^{\infty} \frac{(-\alpha)^n}{n!} t^{2n} \sum_{k=0}^{\infty} (-1)^k c_k^{(n)} t^k = \sum_{m=0}^{\infty} (-1)^m t^m \sum_{n=0}^{[m/2]} \frac{(-\alpha)^n}{n!} c_{m-2n}^{(n)},$$

und dieses zeigt, daß

$$F^{(m)}(0) = (-1)^m m! \sum_{n=0}^{[m/2]} \frac{(-\alpha)^n}{n!} c_{m-2n}^{(n)}. \tag{9}$$

Nun wenden wir uns dem Integral H_2 zu.

Zuerst bemerken wir, um eine obere Schranke für $|F(t)|$ zu erhalten, daß aus der Gleichung

$$\log|F(t)| = \Re\{\alpha[\log(1+t) - t]\},$$

mit einer Derivation nach $|t| = \varrho$, sich

$$\frac{\partial}{\partial \varrho} \log|F(t)| = \Re\left[\alpha\left(\frac{1}{1+t} - 1\right)\frac{\partial t}{\partial \varrho}\right] = -\frac{1}{\varrho}\Re\left[\frac{\alpha t^2}{1+t}\right] = -\frac{\Re[\alpha t^2(1+\bar{t})]}{\varrho|1+t|^2},$$

[4]) Die Formel

$$F^{(m)}(0) = m!\, L_m^{(\alpha-m)}(\alpha),$$

die man von einer bekannten Erzeugungsfunktion der Laguerreschen Polynome leicht ableiten kann, läßt die behauptete Eigenschaft nicht sofort hervortreten.

[5]) In der letzten Zeit habe ich mich überzeugen können, daß öfter, insbesondere bei asymptotischen Untersuchungen, die Koeffizienten der (formellen) n-Potenz einer Potenzreihe nützliche Dienste leisten können. Für Rekurrenzenformeln usw., siehe z. B. K. Knopp, Theorie und Anwendung der unendlichen Reihen (2. Aufl., Berlin, Springer, 1924), S. 178. Die hier vorkommenden Koeffizienten $c_k^{(n)}$ sind mit den Stirlingschen Zahlen verknüpft.

ergibt, d. h.

$$\frac{1}{|F(t)|}\,\frac{\partial\,|F(t)|}{\partial\,\varrho} = -\,\frac{\Re(\alpha\,t^2)+|t|^2\,\Re(\alpha\,t)}{\varrho\,|1+t|^2}.$$

Infolgedessen, solange

$$\Re(\alpha\,t^2)>0, \quad \Re(\alpha\,t)>0$$

ist, d. h. solange

$$\cos(\beta+2\psi)>0, \quad \cos(\beta+\psi)>0, \tag{10}$$

wo β das Argument von α bezeichnet, ist $|F(t)|$ eine abnehmende Funktion von ϱ. Aber für $\varrho=0$ haben wir $F(0)=1$, also unter den obigen Bedingungen ist durchweg

$$|F(t)|\leqq 1.$$

Wenn außer (10) auch (7) befriedigt ist, erlaubt uns die letzte Ungleichung zu schreiben

$$|H_2|\leqq \int_{t_0}^{\infty e^{i\psi}} e^{\Re[(\alpha-x)\,t]}\,|d\,t|,$$

wovon, mit Hilfe der Substitution $t=e^{i\psi}u$,

$$|H_2|\leqq \int_0^{\infty} e^{-u\,|x-\alpha|\,\cos(\xi+\psi)}\,d\,u$$

folgt. Aber für $|x-\alpha|\to\infty$ strebt dieses letzte Integral *exponentiell* gegen Null, weil $\theta>0$, also geschieht dasselbe für H_2 und wir können behaupten, daß unter den Bedingungen (7), (8) und (10) die einfache asymptotische Entwicklung gilt

$$\Gamma(\alpha+1,\,x)\sim e^{-x}\,x^{\alpha+1}\sum_{m=0}^{\infty}\frac{F^{(m)}(0)}{(x-\alpha)^{m+1}}, \tag{11}$$

was insbesondere mit sich zieht, daß

$$\Gamma(\alpha+1,\,x)=\frac{e^{-x}\,x^{\alpha+1}}{x-\alpha}\left[1-\frac{\alpha}{(x-\alpha)^2}+\frac{2\,\alpha}{(x-\alpha)^3}+O(|\alpha|^3\,|x-\alpha|^{-4})\right]. \tag{11'}$$

Endlich müssen wir die Kompatibilität der Bedingungen (7) und (10) diskutieren, wovon die erste unter die Form

$$-\xi-\pi/2<\psi<-\xi+\pi/2$$

gebracht werden kann, während die beiden anderen in die einzige Ungleichung

$$-\beta/2-\pi/4<\psi<-\beta/2+\pi/4$$

zusammengefaßt werden können. (Von Vielfachen von 2π abgesehen).

Damit sehen wir, daß die gesuchten Kompatibilitäts-Bedingungen sind:

$$-\xi-\pi/2<-\beta/2+\pi/4, \quad -\beta/2-\pi/4<-\xi+\pi/2,$$

welche zeigen, daß

$$-\frac{3\pi}{4} < \frac{\beta}{2} - \xi < \frac{3\pi}{4}$$

gelten muß, d. h.

$$\left| \arg \frac{\sqrt{\alpha}}{x-\alpha} \right| < \frac{3\pi}{4}.$$

Wir kommen so zu dem Schluß, daß *für die Gültigkeit der asymptotischen Entwicklung* (11) *hinreichend ist, daß der Modul des Verhältnisses* $z = \sqrt{\alpha}/(x-\alpha)$ *gegen Null strebt, während sein Argument zwischen* $-3\pi/4$ *und* $3\pi/4$ *bleibt.*

Die Bedingung $|\arg z| < 3\pi/4$ kann sicher durch die weniger einschränkende Bedingung $|\arg z| < \pi$ ersetzt werden, aber um den Preis von manchen Komplikationen.

4. Um den *speziellen Fall* zu behandeln, hat sich als nützlich erwiesen, damit anzufangen, das Hilfsintegral

$$I(\alpha, \beta) = \int_0^\beta e^{\alpha t}(1-t)^\alpha \, dt \tag{12}$$

für $|\alpha| \to \infty$ asymptotisch abzuschätzen.

Ähnlich wie in 3. in diesem Falle haben wir

$$e^{\alpha t}(1-t)^\alpha = \exp[\alpha t + \alpha \log(1-t)] = \exp\left[-\frac{\alpha t^2}{2} - \alpha t^3\left(\frac{1}{3} + \frac{t}{4} + \frac{t^2}{5} + \cdots\right)\right]$$

$$= e^{-\alpha t^2/2} \sum_{m=0}^{\infty} \frac{(-\alpha)^m}{m!} t^{3m}\left(\frac{1}{3} + \frac{t}{4} + \frac{t^2}{5} + \cdots\right)^m.$$

Hieraus folgt: wenn wir setzen

$$\left(\frac{1}{3} + \frac{z}{4} + \frac{z^2}{5} + \cdots\right)^m = \sum_{k=0}^{\infty} a_k^{(m)} z^k \qquad (m = 0, 1, 2, \ldots), \tag{13}$$

also [6]) insbesondere

$$a_0^{(m)} = \frac{1}{3^m}, \qquad a_1^{(m)} = \frac{m}{3^{m-1} \cdot 4}, \qquad a_2^{(m)} = \frac{m(15m+17)}{3^{m-1} \cdot 160}, \ldots,$$

so können wir auch schreiben

$$e^{\alpha t}(1-t)^\alpha = e^{-\alpha t^2/2} \sum_{k,\, m=0}^{\infty} \frac{(-\alpha)^m}{m!} a_k^{(m)} t^{3m+k}$$

$$= e^{-\alpha t^2/2} \sum_{n=0}^{\infty} \sum_{m=0}^{n} \frac{(-\alpha)^m}{m!} a_{n-m}^{(m)} t^{n+2m}.$$

---

[6]) Zwischen dem neuen Koeffizienten $a_k^{(m)}$ und dem alten $c_k^{(m)}$ besteht die leicht verifizierbare Beziehung

$$2^m c_{nl}^{(m)} = \sum_{h=0}^{n} \binom{ml}{h} 2^h a_{n-h}^{(h)}.$$

Mit Hilfe der Substitution $t = \beta\sqrt{u}$ bekommen wir

$$\int_0^\beta e^{-\alpha t^2/2} t^{n+2m} dt = \frac{1}{2}\beta^{n+2m+1}\int_0^\beta e^{-\frac{\alpha\beta^2}{2}u} u^{\frac{n-1}{2}+m} du$$

$$= \frac{1}{2}\beta^{n+2m+1}\Gamma\left(\frac{n+1}{2}+m\right)\gamma^*\left(\frac{n+1}{2}+m, \frac{\alpha\beta^2}{2}\right),$$

also, *solange die Funktion*

$$(14)\quad F_{n,m}(\alpha,\beta) = \left(\frac{\alpha}{2}\right)^{\frac{n+1}{2}+m}\beta^{n+1+2m}\Gamma\left(\frac{n+1}{2}+m\right)\gamma^*\left(\frac{n+1}{2}+m, \frac{\alpha\beta^2}{2}\right)$$

$$= \pm\,\gamma\left(\frac{n+1}{2}+m, \frac{\alpha\beta^2}{2}\right)$$

bei wachsendem $|\alpha|$ beschränkt bleibt, gilt die asymptotische Entwicklung

$$(15)\quad I(\alpha,\beta) \sim \frac{1}{2}\sum_{n=0}^{\infty} A_n(\alpha,\beta)\cdot\left(\frac{2}{\alpha}\right)^{\frac{n+1}{2}},$$

wo

$$(16)\quad A_n(\alpha,\beta) = \sum_{m=0}^{n}\frac{(-2)^m}{m!}a_{n-m}^{(m)}F_{n,m}(\alpha,\beta).$$

Mit leichten Änderungen, d. h. mit einer teilweisen Ersetzung von α durch $-\alpha$, läßt sich das vorige Resultat auch in die Gestalt

$$(17)\quad I(\alpha,\beta) \sim \frac{1}{2}\sum_{n=0}^{\infty} A_n^*(\alpha,\beta)\left(-\frac{2}{\alpha}\right)^{\frac{n+1}{2}}$$

bringen, wo

$$(18)\quad A_n^*(\alpha,\beta)$$

$$= \sum_{m=0}^{n}\frac{2^m}{m!}a_{n-m}^{(m)}\left(-\frac{\alpha}{2}\right)^{\frac{n+1}{2}+m}\beta^{n+1+2m}\Gamma\left(\frac{n+1}{2}+m\right)\gamma^*\left(\frac{n+1}{2}+m, \frac{\alpha\beta^2}{2}\right).$$

Es fehlt nur noch die endgültige Bestimmung der Funktionen A_n und A_n^*, wobei man die zwei Fälle:

A) $-\frac{\pi}{2} < \arg\beta \leqq \frac{\pi}{2}$ B) entw. $-\pi < \arg\beta \leqq -\frac{\pi}{2}$ oder $\frac{\pi}{2} < \arg\beta \leqq \pi$

zu unterscheiden sind, weil in diesen Fällen respektive gilt:

A) $\beta^n \cdot \text{H.W.}(\beta^2)^{-n/2} = 1$ B) $\beta^n \cdot \text{H.W.}(\beta^2)^{-n/2} = (-1)^n$.

Mit leichten Rechnungen finden wir so:

$$(19)\quad A_n(\alpha,\beta) = \begin{cases} \displaystyle\sum_{m=0}^{n}\frac{(-2)^m}{m!}a_{n-m}^{(m)}\gamma\left(\frac{n+1}{2}+m, \frac{\alpha\beta^2}{2}\right) & \text{(A)} \\ (-1)^{n+1}\displaystyle\sum_{m=0}^{n}\frac{(-2)^m}{m!}a_{n-m}^{(m)}\gamma\left(\frac{n+1}{2}+m, \frac{\alpha\beta^2}{2}\right) & \text{(B)} \end{cases}$$

$$(20)\qquad A_n^*(\alpha,\beta)=\begin{cases}\displaystyle\sum_{m=0}^{n}\frac{2^m}{m!}\,a_{n-m}^{(m)}\,\gamma_1\Big(\frac{n+1}{2}+m,-\frac{\alpha\beta^2}{2}\Big) & \text{(A)}\\ \displaystyle(-1)^{n+1}\sum_{m=0}^{n}\frac{2^m}{m!}\,a_{n-m}^{(m)}\,\gamma_1\Big(\frac{n+1}{2}+m,-\frac{\alpha\beta^2}{2}\Big) & \text{(B)}\end{cases}$$

Insbesondere hat man

$$(21)\qquad \begin{cases} A_0=\pm\,2\,\mathrm{Erf}\,(\sqrt{\alpha\beta^2/2}), \quad A_0^*=\pm\,2\,\mathrm{Erfi}\,(\sqrt{-\alpha\beta^2/2})\\ A_1=-A_1^*=\frac{2}{3}\Big[e^{-\alpha\beta^2/2}\Big(1+\frac{\alpha\beta^2}{2}\Big)-1\Big]\end{cases}$$

wo das obere (bzw. das untere) Vorzeichen dem Fall (A) bzw. (B) entspricht, und

$$(22)\qquad \mathrm{Erf}\,(z)=\int_0^z e^{-t^2}\,dt=\tfrac{1}{2}\,\gamma(\tfrac{1}{2},z^2),\qquad \mathrm{Erfi}\,(z)=\int_0^z e^{t^2}\,dt=\tfrac{1}{2}\,\gamma_1(\tfrac{1}{2},z^2).$$

5. Nach der vorangehenden Vorbereitung kann der *spezielle Fall*, d. h., genauer präzisiert als vorher, der Fall in welchem bei $\gamma(1+\alpha,x)$ oder $\gamma_1(1-\alpha,x)$ ist [7])

$$(23)\qquad x=\alpha+O(\sqrt{\alpha}),$$

schnell erledigt werden. Man muß nur zwei Unterfälle unterscheiden, jenachdem der Realteil des ursprünglichen Parameters α positiv oder negativ ausfällt, oder (was im Grunde dasselbe ist) die zwei Funktionen $\gamma(1+\alpha,x)$ und $\gamma_1(1-\alpha,x)$ getrennt unter der Annahme $\Re\alpha>0$ betrachtet werden, was wir hier tun wollen.

Im ersten Unterfall erhalten wir aus (1), mit Hilfe der Substitution $t=\alpha(1-\tau)$ unter Benutzung der Position (12),

$$\begin{aligned}\gamma(\alpha+1,\alpha+\xi)&=e^{-\alpha}\alpha^{\alpha+1}\int_{-\xi/\alpha}^{1}e^{\alpha\tau}(1-\tau)^{\alpha}\,d\tau\\ &=e^{-\alpha}\alpha^{\alpha+1}\,[I(\alpha,1)-I(\alpha,-\xi/\alpha)]\end{aligned}$$

Aber für $\beta=1$ haben wir

$$\lim_{\alpha\to\infty}F_{m,n}(\alpha,\beta)=\pm\lim_{n\to\infty}\gamma\Big(\frac{n+1}{2}+m,\frac{\alpha\beta^2}{2}\Big)=\pm\,\Gamma\Big(\frac{n+1}{2}+m\Big)$$

während für $\beta=-\xi/\alpha$ dieser Grenzwert ebenfalls endlich ist, nämlich gleich

$$0,\ \pm\,\gamma\Big(\frac{n+1}{2}+m,k\Big),\qquad \pm\,\Gamma\Big(\frac{n+1}{2}+m\Big)$$

je nachdem

$$\xi=o(\sqrt{\alpha}),\qquad \lim(\xi\alpha^{-1/2})=k,\qquad \sqrt{\alpha}=o(\xi)$$

[7]) Diese Bedingung ist aber nur in Nr. 6 wesentlich.

also ist die Bedingung betreffs der Beschränktheit der Funktion $F_{m,n}$ ebenso für $\beta = 1$ wie für $\beta = -\xi/\alpha$ befriedigt, und aus (15) und aus der Bemerkung, daß der Unterschied zwischen

$$\gamma\left(\frac{n+1}{2}+m, \frac{\alpha}{2}\right) \quad \text{und} \quad \Gamma\left(\frac{n+1}{2}+m\right)$$

für $\alpha \to \infty$ *exponentiell* verschwindet, bekommen wir ohne weiteres:

$$(24)\quad \boxed{\gamma(\alpha+1,\ \alpha+\sqrt{2\alpha}\,y) \sim \sqrt{\frac{\alpha}{2}}\, e^{-\alpha}\alpha^{\alpha} \sum_{n=0}^{\infty} B_n(y)\cdot\left(\frac{2}{\alpha}\right)^{n/2} \quad (\Re\,\alpha > 0,\ \alpha \to \infty)}$$

wo

$$(25)\quad B_n(y) = \begin{cases} \displaystyle\sum_{m=0}^{n} \frac{(-2)^m}{m!}\, a_{n-m}^{(m)}\, \Gamma\left(\frac{n+1}{2}+m, y^2\right) \\ \qquad\qquad \left(\frac{\pi}{2} < |\arg y - \tfrac{1}{2}\arg\alpha| < \pi\right) \\ \displaystyle\sum_{m=0}^{n} \frac{(-2)^m}{m!}\, a_{n-m}^{(m)} \left[\Gamma\left(\frac{n+1}{2}+m\right) + (-1)^m \gamma\left(\frac{n+1}{2}+m, y^2\right)\right] \\ \qquad\qquad \left(|\arg y - \tfrac{1}{2}\arg\alpha| < \frac{\pi}{2}\right) \end{cases}$$

Vernachlässigen wir die Glieder mit $n \geqq 2$ und beachten außerdem, daß

$$\Gamma(\alpha+1) = \sqrt{2\pi\alpha}\, e^{-\alpha}\alpha^{\alpha}[1 + O(\alpha^{-1})]$$

dann können wir in einfacherer Weise schreiben, daß

$$(26)\quad \frac{\gamma(\alpha+1,\ \alpha+\sqrt{2\alpha}\,y)}{\Gamma(\alpha+1)} = \frac{1}{2} \mp \operatorname{Erf}(\sqrt{y^2}) - \frac{1}{3}\sqrt{\frac{2}{\alpha\pi}}\,(1+y^2)\,e^{-y^2} + O\left(\frac{1}{\alpha}\right)$$

wo das obere (bzw. das untere) Vorzeichen dem ersten (bzw. dem zweiten) Falle von (25) entspricht.

Noch spezieller: wenn α und y beide *reell* sind, dann können wir (ohne den Fall $y > 0$ von dem Falle $y < 0$ unterscheiden zu müssen) schreiben:

$$(27)\quad \frac{\gamma(\alpha+1,\ \alpha+\sqrt{2\alpha}\,y)}{\Gamma(\alpha+1)} = \frac{1}{2} + \frac{1}{\sqrt{\pi}} \operatorname{Erf}(y) - \frac{1}{3}\sqrt{\frac{2}{\alpha\pi}}\,(1+y^2)\,e^{-y^2} + O\left(\frac{1}{\alpha}\right).$$

Unter anderem erlaubt uns diese einfache und (vom numerischen Standpunkt aus) sehr genaue asymptotische Formel, eine gute Abschätzung der Teilsummen der Exponentialreihe:

$$e_n(x) \equiv \sum_{h=0}^{n} \frac{x^h}{h!} = e^x\left[1 - \frac{\gamma(n+1, x)}{n!}\right],$$

welche in manchen Fragen der modernen Physik (z. B. Theorie der Geigerzähler) Interesse bekommen haben, in der Nähe des kritischen

Wertes $x = n$ zu gewinnen. Wir finden nämlich, daß

(28)
$$e_n(n + \sqrt{2n}\,y) = e^{n+\sqrt{2n}\,y}\left[\frac{1}{2} - \frac{1}{\sqrt{\pi}}\,\mathrm{Erf}\,(y) + \frac{1}{3}\sqrt{\frac{2}{n\pi}}\,(1+y^2)\,e^{-y^2} + O\left(\frac{1}{n}\right)\right].$$

6. *Im zweiten Unterfalle* kann man Integraldarstellungen des Typus (1) nicht direkt benutzen wegen der Singularität des Integranden im Ursprung, aber es dürfen Integraldarstellungen dieser Art für die Berechnung der Unterschiede zwischen zwei verschiedenen Werten der Funktion $\gamma(\alpha, x)$ oder $\gamma_1(\alpha, x)$ verwendet werden. Insbesondere können wir schreiben, daß

$$\gamma_1(1-\alpha,\ \alpha+\xi) - \gamma_1(1-\alpha,\ \alpha) = \int_{\alpha}^{\alpha+\xi} e^t\, t^{-\alpha}\, dt$$
$$= -e^{\alpha}\,\alpha^{1-\alpha}\int_{0}^{-\xi/\alpha} e^{-\alpha\tau}\,(1-\tau)^{-\alpha}\,d\tau = -e^{\alpha}\,\alpha^{1-\alpha}\,I(-\alpha,\ -\xi/\alpha).$$

Aber, solange $\xi = O(\sqrt{\alpha})$, bleibt die Funktion (14), d. h. die Funktion

$$F_{n,m}(-\alpha, -\xi/\alpha) = \pm\,\gamma\left(\frac{n+1}{2} + m,\ -\frac{\xi^2}{2\alpha}\right)$$

offenkundig beschränkt bei wachsendem $|\alpha|$, und wir erhalten unter der obigen Einschränkung

$$\gamma_1(1-\alpha,\ \alpha+\sqrt{2\alpha}\,y) - \gamma_1(1-\alpha,\ \alpha) \sim -\sqrt{\frac{\alpha}{2}}\,e^{\alpha}\,\alpha^{-\alpha}\sum_{n=0}^{\infty} A_n^*(y)\cdot\left(\frac{2}{\alpha}\right)^{n/2} \tag{29}$$
$$(\Re\,\alpha > 0),$$

wo

$$A_n^*(y) = \begin{cases} \displaystyle\sum_{m=0}^{n} \frac{2^m}{m!}\,a_{n-m}^{(m)}\,\gamma_1\left(\frac{n+1}{2}+m,\ y^2\right) \\ \qquad\qquad \left(\frac{\pi}{2} < \left|\arg y - \frac{1}{2}\arg\alpha\right| < \pi\right), \\ (-1)^{n+1}\displaystyle\sum_{m=0}^{n} \frac{2^m}{m!}\,a_{n-m}^{(m)}\,\gamma_1\left(\frac{n+1}{2}+m,\ y^2\right) \\ \qquad\qquad \left(\left|\arg y - \frac{1}{2}\arg\alpha\right| < \frac{\pi}{2}\right). \end{cases} \tag{30}$$

Sind α und y beide reell und werden die Glieder mit $n \geqq 2$ vernachlässigt, dann haben wir einfacher:

$$\gamma_1(1-\alpha,\ \alpha+\sqrt{2\alpha}\,y) - \gamma_1(1-\alpha,\ \alpha)$$
$$= \sqrt{2\alpha}\,e^{\alpha}\,\alpha^{-\alpha}\left\{\mathrm{Erfi}\,(y) - \frac{1}{3}\sqrt{\frac{2}{\alpha}}\left[1-(1-y^2)\,e^{y^2}\right] + O\left(\frac{1}{\alpha}\right)\right\} \tag{31}$$
$$(\Re\,\alpha > 0,\ \alpha \to \infty,\ y = O(1)).$$

Um unsere Untersuchung zu vervollkommnen, müssen wir jetzt die Grösse $\gamma_1(1-\alpha,\ \alpha)$ asymptotisch abschätzen, wobei uns eine vermutlich

bisher unbeachtet gebliebene Integraldarstellung der unvollständigen Gammafunktion zustatten kommt, nämlich die Formel

$$\gamma(\alpha, x) = e^{-x} x^{\alpha} [F_1 \cdot \operatorname{cotg}(\alpha\pi) + F_2], \left(\Re\alpha \leqq 0, \left|\arg(-x)\right| \leqq \frac{\pi}{2}\right) \tag{32}$$

wo man, unter der Annahme $x = -\xi e^{\varphi i}$,

$$\begin{aligned} F_1 &= \frac{1}{2} e^{-\varphi i} \int_{-\infty}^{\infty} e^{i\xi t}(1 + i e^{-\varphi i} t)^{\alpha - 1} dt, \\ F_2 &= \frac{1}{2i} e^{-\varphi i} \int_{-\infty}^{\infty} e^{i\xi t}(1 + i e^{-\varphi i} t)^{\alpha - 1} \operatorname{sgn}(t)\, dt \end{aligned} \tag{33}$$

gesetzt hat. In einfacherer Weise, wenn α *und* x *beide reell sind*, hat man

$$\gamma(\alpha, x) = \frac{e^{-x} x^{\alpha}}{\sin(\alpha\pi)} \Re\left[e^{-\alpha\pi i} \int_0^{\infty} e^{-ixt}(1 + it)^{\alpha - 1} dt\right] \tag{34}$$

mit der Bemerkung, daß die Formel sogar für $\alpha = 0, -1, -2, \ldots$ gültig ist, wenn $\gamma(\alpha, x)$ durch $\gamma^*(\alpha, x)$ ausgedrückt wird.

Wenn wir — der Einfachheit halber — die Annahme der Realität von α beibehalten, ergibt die letzte Formel

$$\begin{aligned} \gamma_1(1-\alpha, \alpha) &= -\frac{e^{\alpha} \alpha^{1-\alpha}}{\sin(\alpha\pi)} \Re\left[e^{\alpha\pi i} \int_0^{\infty} e^{ixt}(1 + it)^{-\alpha} dt\right] \\ &= -\frac{e^{\alpha} \alpha^{1-\alpha}}{\sin(\alpha\pi)} \Re\left[i e^{\alpha\pi i} \int_0^{-i\infty} e^{-\alpha\tau}(1 - \tau)^{-\alpha} d\tau\right] \\ &= -\frac{e^{\alpha} \alpha^{1-\alpha}}{\sin(\alpha\pi)} \Re[i\, l(-\alpha, -i\infty)]. \end{aligned}$$

Da die Bedingung betreffs der Funktion (14) auch jetzt befriedigt ist, bekommen wir nun aus (17) und (20):

$$i\, l(-\alpha, -i\infty) \sim \frac{1}{\sqrt{2\alpha}} \sum_{n=0}^{\infty} \left(-i\sqrt{\frac{2}{\alpha}}\right)^n \sum_{m=0}^{n} \frac{(-2)^m}{m!} a_{n-m}^{(m)} \Gamma\left(\frac{n+1}{2} + m\right)$$

und mit der Abkürzung

$$K_n = \sum_{m=0}^{n} \frac{(-2)^m}{m!} a_{n-m}^{(m)} \Gamma\left(\frac{n+1}{2} + m\right) \quad (n = 0, 1, 2, \ldots) \tag{35}$$

erhalten wir daraus die asymptotische Entwicklung

$$\begin{aligned} \gamma_1(1-\alpha, \alpha) \sim -\sqrt{\frac{\alpha}{2}}\, e^{\alpha} \alpha^{-\alpha} \Bigg[&\operatorname{cotg}(\alpha\pi) \sum_{n=0}^{\infty} K_{2n} \cdot \left(-\frac{2}{\alpha}\right)^n \\ &+ \sqrt{\frac{2}{\alpha}} \sum_{n=0}^{\infty} K_{2n+1} \cdot \left(-\frac{2}{\alpha}\right)^n\Bigg] \end{aligned} \tag{36}$$

$$(\alpha \to +\infty).$$

Insbesondere haben wir

$$\gamma_1(1-\alpha,\alpha) = -\sqrt{2\alpha}\,e^{\alpha}\,\alpha^{-\alpha}\left[\sqrt{\frac{\pi}{2}}\operatorname{cotg}(\alpha\pi) - \frac{1}{3}\sqrt{\frac{2}{\alpha}} + O\left(\frac{1}{\alpha}\right)\right] \tag{37}$$

und, nach Addition von (31):

$$\gamma_1(1-\alpha,\alpha+\sqrt{2\alpha}\,y) = \sqrt{2\alpha}\,e^{\alpha}\,\alpha^{-\alpha}\left[-\sqrt{\frac{\pi}{2}}\operatorname{cotg}(\alpha\pi) - \operatorname{Erfi}(y) + \frac{1}{3}\sqrt{\frac{2}{\alpha}}\,(1-y^2)\,e^{y^2} + O\left(\frac{1}{\alpha}\right)\right]. \tag{38}$$

Da

$$\Gamma(\alpha) = \sqrt{\frac{2\pi}{\alpha}}\,e^{-\alpha}\,\alpha^{\alpha}\,[1+O(\alpha^{-1})]$$

können wir also auch schreiben:

$$\boxed{\begin{gathered}\Gamma(\alpha)\,\gamma_1(1-\alpha,\alpha+\sqrt{2\alpha}\,y) = -\pi\operatorname{cotg}(\alpha\pi) + 2\sqrt{\pi}\,\operatorname{Erfi}(y) \\ + \frac{2}{3}\sqrt{\frac{2\pi}{\alpha}}\,(1-y^2)\,e^{y^2} + O\left(\frac{1}{\alpha}\right) \\ \alpha, y \text{ reell},\ \alpha>0;\quad \alpha\to+\infty,\ y=0(1)\end{gathered}} \tag{39}$$

7. Die allgemeinen Ergebnisse über die reellen Nullstellen der konfluenten hypergeometrischen Funktionen[8]) erlauben uns, die reellen Nullstellen der Funktion $\gamma^*(\alpha,x)$[9]), bei reellen α, sofort abzuzählen. Man findet nämlich, daß es solange $\alpha>0$ oder $\alpha=0,-1,-2,\ldots$ ist, *keine* (abgesehen von $x=0$ und $x=\infty$) solche Nullstelle gibt, aber für $\alpha<0$ gibt es entweder *eine negative Nullstelle* oder *eine negative und eine positive Nullstelle*, je nachdem der *ganze Teil* von $-\alpha$ *gerade* oder *ungerade* ist.

Diese Nullstellen scheinen bis jetzt gar nicht beachtet worden zu sein, vielleicht weil bis jetzt der Parameter α fast immer stillschweigend als positiv betrachtet worden war.

In meinem im Anfang erwähnten Bericht über die unvollständigen Gammafunktionen, werden u. a. asymptotische Ausdrücke für $|\alpha|\to\infty$ ebenso für die negative Nullstelle x' wie für die positive Nullstelle x'' unserer Funktion aus den vorangehenden asymptotischen Formeln (39) und (11) abgeleitet. So findet man für x' die einfache asymptotische Darstellung

$$x' = -(1-\alpha)\left[1+\sqrt{\frac{2}{1-\alpha}}\,y^*(\alpha) + O\left(\frac{1}{\alpha}\right)\right], \tag{40}$$

[8]) Siehe z. B. meine Arbeit: *Über die Abzählung der Nullstellen der konfluenten hypergeometrischen Funktionen* in derselben Zeitschrift.

[9]) Welche, abgesehen von $x=0$ und $x=\infty$, mit den Nullstellen von $\gamma(\alpha,x)$ oder $\gamma_1(\alpha_1,-x)$ zusammenfallen.

wo $y^*(\alpha)$ die einzige positive Wurzel der transzendenten Gleichung

$$\text{Erfi}\,(y) = \sqrt{\frac{\pi}{2}}\,\text{cotg}\,(\alpha y) \tag{41}$$

bezeichnet. Dagegen findet man unter der Annahme, daß $[-\alpha]$ gerade sei, d. h. unter der Annahme

$$\alpha < 0, \quad \sin(\alpha\pi) > 0$$

für x'' den asymptotischen Ausdruck

$$x'' = -\tau\alpha - \frac{\tau}{1+\tau}\log\left[\frac{1+\tau\sqrt{-\alpha\pi/2}}{\sin(\alpha\pi)}\right] + O\left\{\frac{[\log(-\alpha)]^2}{\alpha}\right\}, \tag{42}$$

wo τ — eine Weltkonstante — die einzige positive Wurzel der transzendenten Gleichung

$$1 + x + \log x = 0 \tag{43}$$

bedeutet. Man findet ungefähr $\tau = 0{,}278\,463$.

8. Zum Schluß will ich hier einige elegante und teilweise recht gut konvergierende Reihenentwicklungen für die unvollständigen Gammafunktionen und ihren Sonderfall $\alpha = \frac{1}{2}$ (Fehlerintegral) erwähnen, die neu oder wenigstens wenig bekannt zu sein scheinen. Ohne Beweise aber, um diese schon zu lange Note nicht ungebührlich zu verlängern.

$$\Gamma(\alpha, x) = e^{-x}x^{\alpha}\sum_{n=0}^{\infty}\frac{L_n^{(\alpha)}(x)}{n+1} \quad (x > 0,\ L_n^{(\alpha)} \equiv \text{Laguerresches Polynom}), \tag{44}$$

$$\gamma(\alpha, \lambda x) = \lambda^{\alpha}\sum_{n=0}^{\infty}\gamma(\alpha+n, x)\frac{(1-\lambda)^n}{n!} \quad \text{(unbedingt gültig)}, \tag{45}$$

$$\gamma(\alpha, x) = \Gamma(\alpha)\,e^{-x}x^{\alpha/2}\sum_{n=0}^{\infty}e_n(-1)\,x^{n/2}I_{n+\alpha}(2\sqrt{x}) \tag{46}$$

$$(\alpha > -\tfrac{1}{2},\ I_\nu \equiv \text{mod. Zylinderfunktion}),$$

$$\text{Erf}\,(\sqrt{x+y}) = \text{Erf}\,(\sqrt{x}) + \frac{e^{-x}}{2\sqrt{x}}\sum_{n=0}^{\infty}(-1)^n\frac{1\cdot 3\cdot\ldots\cdot(2n-1)}{2\cdot 4\cdot\ldots\cdot(2n)}\,\frac{\gamma(n+1, y)}{x^{n+1}} \tag{47}$$

$$(|y| < |x|),$$

$$\text{Erf}\,(x) = \tfrac{1}{2}\sqrt{\pi x}\,e^{-x^2}\sum_{n=0}^{\infty}e_n(-1)\,x^n I_{n+1/2}(2x) \tag{48}$$

$$\text{(unb. gültig, } I_\nu \text{ wie oben)},$$

$$\text{Erf}\,(\sqrt{x}) = \sqrt{\frac{\pi}{2}}\sum_{n=1}^{\infty}(-1)^{[n/2]}I_{n-1/2}(x) \tag{49}$$

$$\text{Erfi}\,(\sqrt{x}) = \sqrt{\frac{\pi}{2}}\sum_{n=0}^{\infty}(-1)^{[n/2]}I_{n+1/2}(x) \tag{50}$$

(unbedingt gültig, $[n/2]$ = ganze Teil von $n/2$).

Vielleicht können manche von diesen Entwicklungen auch bei numerischen Rechnungen gute Dienste leisten.

(Eingegangen am 3. Januar 1950.)

Mathematische Zeitschrift, Band 53, Heft 2, S. 149—161 (1950).

Über die Konvergenzkriterien bei Iterationsverfahren für lineare Gleichungssysteme.

Herrn Erich Kamke zum 60. Geburtstag am 18. August 1950 gewidmet.

Von

L. Collatz in Hannover.

Für verschiedene Arten von Iterationsverfahren für lineare Gleichungssysteme, von denen die Verfahren „in Gesamtschritten" und „in Einzelschritten" wohl die am meisten benutzten sind, wurde eine Reihe von Konvergenzkriterien, meist mittels elementarer Methoden aufgestellt; im folgenden soll gezeigt werden, daß diese Kriterien[1]) durch Anwendung einfacher Sätze aus der Matrizenlehre einheitlich und kurz hergeleitet werden können. Zugleich werden einige Kriterien etwas erweitert und auf eine für die Anwendungen besonders bequeme Form gebracht.

1. Notwendige und hinreichende Konvergenzbedingung.

Im folgenden bezeichnen große deutsche Buchstaben $\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \ldots$ quadratische n-reihige Matrizen mit den Elementen $a_{jk}, b_{jk}, c_{jk}, \ldots$ und kleine deutsche Buchstaben $\mathfrak{x}, \mathfrak{y}, \mathfrak{r}, \ldots$ einspaltige Matrizen oder „Vektoren" mit den Elementen $x_j, y_j, r_j, \ldots$; $\mathfrak{E}$ ist die n-reihige Einheitsmatrix mit den Elementen $e_{jk} = \begin{cases} 0 & \text{für } j \neq k \\ 1 & \text{für } j = k \end{cases}$, $\det \mathfrak{A}$ die Determinante

[1]) R. v. Mises und H. Pollaczek-Geiringer, Praktische Verfahren der Gleichungsauflösung, Z. angew. Math. Mech. 9 (1929), S. 58—77, 152—164.

H. Wittmeyer, Einfluß der Änderung einer Matrix auf die Lösung des zugehörigen Gleichungssystems, sowie auf die charakteristischen Zahlen und die Eigenvektoren, Z. angew. Math. Mech. 16 (1936), S. 287—300.

H. Wittmeyer, Über die Lösung von linearen Gleichungssystemen durch Iteration, Z. angew. Math. Mech. 16 (1936), S. 301—310.

L. Cesari, Sulla risoluzioni dei sistemi di equazioni lineari per approssimazioni successive. Rendic. Reale Accademia Nazionale dei Lincei, Classe Scienze fis. math. natur., 25 (ser. 6a) Roma 1937, S. 422—428.

L. Collatz, Fehlerabschätzung für das Iterationsverfahren zur Auflösung linearer Gleichungssysteme, Z. angew. Math. Mech. 22 (1942), S. 357—361.

E. Bodewig, Bericht über die verschiedenen Methoden zur Lösung eines Systems linearer Gleichungen mit reellen Koeffizienten I—V, Konifiklijke nederlandsche Akademie van Wetenschappen 1947, S. 441—452, 1104—1116, 1285—1295, 1948, S. 53—64, 211—219.

von $\mathfrak{A}$. Die Elemente dürfen reelle oder komplexe Zahlen sein. Die charakteristischen Zahlen $\varkappa$ einer Matrix $\mathfrak{A}$ sind die Wurzeln von $\det(\mathfrak{A} - \varkappa\mathfrak{E}) = 0$.

Vorgelegt sei das lineare Gleichungssystem

$$\mathfrak{A}\mathfrak{x} = \mathfrak{r}, \tag{1}$$

wobei $\mathfrak{r}$ ein gegebener Vektor und die Koeffizientenmatrix $\mathfrak{A}$ nichtsingulär ist, d. h. $\det \mathfrak{A} \neq 0$. $\mathfrak{A}$ werde als Summe zweier Matrizen $\mathfrak{B}$ und $\mathfrak{C}$ geschrieben:

$$\mathfrak{A} = \mathfrak{B} + \mathfrak{C}; \quad \mathfrak{B}\mathfrak{x} + \mathfrak{C}\mathfrak{x} = \mathfrak{r}. \tag{2}$$

Ausgehend von einem willkürlich gewählten Vektor $\mathfrak{x}^{(0)}$ wird beim Iterationsverfahren eine Folge von Vektoren $\mathfrak{x}^{(1)}, \mathfrak{x}^{(2)}, \ldots$ nach der Vorschrift bestimmt

$$\mathfrak{B}\mathfrak{x}^{(\nu+1)} + \mathfrak{C}\mathfrak{x}^{(\nu)} = \mathfrak{r} \qquad (\nu = 0, 1, 2, \ldots). \tag{3}$$

Damit $\mathfrak{x}^{(\nu+1)}$ eindeutig berechnet werden kann, wird $\mathfrak{B}$ als nichtsingulär vorausgesetzt: $\det \mathfrak{B} \neq 0$; es existiert dann die Matrix $\mathfrak{F} = -\mathfrak{B}^{-1}\mathfrak{C}$. Für die „Änderungsvektoren"

$$\mathfrak{d}^{(\nu)} = \mathfrak{x}^{(\nu+1)} - \mathfrak{x}^{(\nu)} \qquad (\nu = 0, 1, 2, \ldots) \tag{4}$$

gilt dann mit $\mathfrak{B}\mathfrak{d}^{(\nu+1)} + \mathfrak{C}\mathfrak{d}^{(\nu)} = 0$

$$\mathfrak{d}^{(\nu+1)} = \mathfrak{F}\mathfrak{d}^{(\nu)} = \cdots = \mathfrak{F}^{\nu+1}\mathfrak{d}^{(0)} \qquad (\nu = 0, 1, 2, \ldots) \tag{5}$$

und

$$\mathfrak{x}^{(\nu+1)} = \mathfrak{x}^{(0)} + \sum_{\varrho=0}^{\nu} \mathfrak{d}^{(\varrho)} = \mathfrak{x}^{(0)} + \left(\sum_{\varrho=0}^{\nu} \mathfrak{F}^{\varrho}\right)\mathfrak{d}^{(0)} \qquad (\nu = 0, 1, 2, \ldots). \tag{6}$$

Nach bekannten Sätzen[2]) konvergiert[3]) für $\nu \to \infty$ die rechtsstehende Summe für jeden beliebigen Vektor $\mathfrak{d}^{(0)}$ genau dann, wenn alle charakteristischen Zahlen $\varkappa$ der Matrix $\mathfrak{F}$, also die Wurzeln von

$$\det(-\mathfrak{B}^{-1}\mathfrak{C} - \varkappa\mathfrak{E}) = \det(-\mathfrak{B}^{-1}) \det(\varkappa\mathfrak{B} + \mathfrak{C}) = 0, \tag{7}$$

Beträge kleiner als 1 haben. Wegen

$$\mathfrak{B}\mathfrak{d}^{(0)} + \mathfrak{A}\mathfrak{x}^{(0)} = \mathfrak{r} \tag{8}$$

entspricht jedem $\mathfrak{d}^{(0)}$ genau ein $\mathfrak{x}^{(0)}$ und umgekehrt; es ist daher die willkürliche Wahl von $\mathfrak{x}^{(0)}$ mit der von $\mathfrak{d}^{(0)}$ gleichwertig; somit gilt:

---

[2]) Vgl. etwa W. Schmeidler, Vorträge über Determinanten und Matrizen, Bln. 1949.

R. Courant-D. Hilbert, Methoden der mathematischen Physik, Bd. I, 2. Aufl., Bln. 1931.

L. Collatz, Eigenwertaufgaben mit technischen Anwendungen, Lpz. 1949, 6. Kapitel, insbesondere S. 281, 311.

[3]) Ähnlich bei Cesari, a. a. O.

Das Iterationsverfahren (2) (3) *zur Auflösung des linearen Gleichungssystems* (1) *mit nichtsingulären Matrizen* $\mathfrak{A}$ *und* $\mathfrak{B}$ *ist bei beliebigem Ausgangsvektor* $\mathfrak{x}^{(0)}$ *genau dann konvergent, wenn alle Wurzeln* $\varkappa$ *von*

$$\det(\varkappa\mathfrak{B} + \mathfrak{C}) = 0 \tag{9}$$

Beträge kleiner als 1 haben, und es ist $\lim_{\nu\to\infty} \mathfrak{x}^{(\nu)} = \mathfrak{x}$.

Für die Fehlervektoren

$$\mathfrak{f}^{(\nu)} = \mathfrak{x}^{(\nu)} - \mathfrak{x}; \qquad (\nu = 0, 1, 2, \ldots)$$

gilt die gleiche Beziehung wie für die $\mathfrak{d}^{(\nu)}$; es folgt somit

$$\mathfrak{B}\mathfrak{f}^{(\nu+1)} + \mathfrak{C}\mathfrak{f}^{(\nu)} = 0$$

$$\mathfrak{f}^{(\nu+1)} = \mathfrak{F}^{\nu+1}\mathfrak{f}^{(0)} \quad \text{und} \quad \mathfrak{F}\mathfrak{d}^{(\nu)} = (\mathfrak{F} - \mathfrak{E})\mathfrak{f}^{(\nu+1)} \quad (\nu = 0, 1, 2, \ldots). \tag{10}$$

Nun setzen wir voraus, daß keine charakteristische Zahl $\varkappa_{\mathfrak{F}}$ der Matrix $\mathfrak{F}$ gleich 1 ist; (wenn z. B. das Iterationsverfahren bei beliebigem $\mathfrak{x}^{(0)}$ konvergiert, ist $|\varkappa_{\mathfrak{F}}| < 1$, also sicher $\varkappa_{\mathfrak{F}} \neq 1$). Dann sind alle charakteristischen Zahlen von $\mathfrak{F} - \mathfrak{E}$ von Null verschieden, und $|\mathfrak{F} - \mathfrak{E}| \neq 0$ [4]). Das ergibt die Fehlerabschätzung

$$|\mathfrak{f}^{(\nu+1)}| \leqq \frac{|\mathfrak{F}|\,|\mathfrak{d}^{(\nu)}|}{|\mathfrak{F} - \mathfrak{E}|} \qquad (\nu = 0, 1, 2, \ldots). \tag{11}$$

Wir setzen nun weiter voraus, daß alle Wurzeln $\varkappa$ von (9) dem Betrage nach kleiner als 1 sind. Dann gibt es eine Zahl α mit $|\varkappa| < \alpha < 1$. Multipliziert man die zweite Gleichung (10) mit $\mathfrak{E}, \mathfrak{F}, \mathfrak{F}', \ldots, \mathfrak{F}^{k-1}$ und addiert, so folgt:

$$\mathfrak{f}^{(\nu+1)} = \mathfrak{F}^k\mathfrak{f}^{(\nu+1)} - \left(\sum_{\varrho=1}^{k} \mathfrak{F}^{\varrho}\right)\mathfrak{d}^{(\nu)} \qquad (\nu = 0, 1, \ldots;\ k = 1, 2, \ldots). \tag{12}$$

Nach unseren Voraussetzungen konvergiert $\mathfrak{F}^k$ für $k \to \infty$ gegen Null, es konvergiert $\sum_{\varrho=0}^{\infty} \mathfrak{F}^{\varrho}$ absolut, und es gibt eine Konstante F, so daß für die Elemente $f_{jk}^{(\varrho)}$ von $\mathfrak{F}^{\varrho}$ gilt

$$|f_{jk}^{(\varrho)}| < F\,\alpha^{\varrho}. \tag{13}$$

Damit wird

$$\sum_{\varrho=1}^{\infty} |f_{jk}^{(\varrho)}| < \sum_{\varrho=1}^{\infty} F \cdot \alpha^{\varrho} = \frac{\alpha}{1-\alpha} F. \tag{14}$$

---

[4]) Wir verwenden hier eine Bezeichnungsweise von H. Wielandt: Hat bei gegebener Matrix $\mathfrak{A}$ die Matrix $\mathfrak{B} = \overline{\mathfrak{A}}'\mathfrak{A}$ die charakteristischen Zahlen $\varkappa_{\mathfrak{B}}$, so ist $|\mathfrak{A}| = +\sqrt{(\varkappa_{\mathfrak{B}})_{\max}}$, $|\mathfrak{A}| = +\sqrt{(\varkappa_{\mathfrak{B}})_{\min}}$. Es gilt dann für beliebige Vektoren $\mathfrak{y}$

$$\cdot\ |\mathfrak{A}|\,|\mathfrak{y}| \leqq |\mathfrak{A}\mathfrak{y}| < |\mathfrak{A}|\,|\mathfrak{y}|.$$

Für $k \to \infty$ folgt dann aus (12) für die Komponenten $x_j^{(\nu)}$ von $\mathfrak{x}^{(\nu)}$ die Fehlerabschätzung

$$(15)\quad |x_j^{(\nu+1)} - x_j| < \frac{\alpha}{1-\alpha}\, nF \operatorname*{Max}_k |x_k^{(\nu+1)} - x_k^{(\nu)}| \quad (j = 1, \ldots, n;\ \nu = 0, 1, \ldots).$$

Je nach der Zerlegung (2) von $\mathfrak{A}$ in $\mathfrak{B} + \mathfrak{C}$ kann man verschiedene Iterationsverfahren als Spezialfälle erhalten. Für

$$b_{jk} = \begin{cases} a_{jk} & \text{für } j = k \\ 0 & \text{für } j \neq k \end{cases}, \qquad c_{jk} = \begin{cases} 0 & \text{für } j = k \\ a_{jk} & \text{für } j \neq k \end{cases}$$

hat man das Verfahren „in Gesamtschritten", für

$$b_{jk} = \begin{cases} a_{jk} & \text{für } j \geqq k \\ 0 & \text{für } j < k \end{cases}, \qquad c_{jk} = \begin{cases} 0 & \text{für } j \geqq k \\ a_{jk} & \text{für } j < k \end{cases}$$

das Verfahren „in Einzelschritten". Bei diesen Verfahren müssen alle $a_{jj} \neq 0$ sein, damit $\mathfrak{B}$ nichtsingulär ist. Für

$$\mathfrak{B} = -\frac{1}{c}\mathfrak{E}, \qquad \mathfrak{C} = \mathfrak{A} + \frac{1}{c}\mathfrak{E}$$

bekommt man ein bei v. Mises, Pollaczek-Geiringer, a. a. O., S. 68 und allgemeiner für $\mathfrak{B} = -\mathfrak{D}^{-1}$, $\mathfrak{C} = \mathfrak{A} + \mathfrak{D}^{-1}$ ein bei Wittmeyer, a. a. O., S. 302 untersuchtes Verfahren, wobei $\mathfrak{D}$ eine beliebige nichtsinguläre Matrix ist[5]).

Für die ersten beiden Spezialfälle gilt dann: Bei nichtsingulärem $\mathfrak{A}$ und $a_{jj} \neq 0$ ist für die Konvergenz des Gesamtschrittverfahrens bei beliebigem Ausgangsvektor $\mathfrak{x}^{(0)}$ notwendig und hinreichend, daß alle Wurzeln $\varkappa$ von

$$(16)\quad \begin{vmatrix} \varkappa a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ a_{21} & \varkappa a_{22} & a_{23} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & a_{n3} & \ldots & \varkappa a_{nn} \end{vmatrix} = 0$$

Beträge kleiner als 1 haben, und für die Konvergenz des Einzelschrittverfahrens, daß alle Wurzeln von

$$(17)\quad \begin{vmatrix} \varkappa a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ \varkappa a_{21} & \varkappa a_{22} & a_{23} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \varkappa a_{n1} & \varkappa a_{n2} & \varkappa a_{n3} & \ldots & \varkappa a_{nn} \end{vmatrix} = 0$$

Beträge kleiner als 1 haben.

[5]) Der obige Satz liefert für dieses Verfahren: Die bei Wittmeyer aufgestellte hinreichende Bedingung, daß alle charakteristischen Zahlen der Matrix $\mathfrak{E} + \mathfrak{D}\mathfrak{A}$ Beträge kleiner als 1 haben, ist auch notwendig dafür, daß das Verfahren für jeden Ausgangsvektor $\mathfrak{x}^{(0)}$ konvergiert.

Sind alle $a_{jj} \neq 0$, so kann man die j-te Gleichung in (1) durch a_{jj} durchdividieren. Es entsteht dann das „durchdividierte Gleichungssystem", bei dem in der Hauptdiagonale der neuen Koeffizientenmatrix $\mathfrak{A}^*$ alle n Zahlen gleich 1 sind. Gl. (16) ist dann die charakteristische Gleichung für die Matrix $\mathfrak{E} - \mathfrak{A}^*$.

2. Positiv definite Matrizen.

Jetzt sei die Koeffizientenmatrix $\mathfrak{A}$ positiv definit, d. h. sie sei hermitesch, $\mathfrak{A} = \overline{\mathfrak{A}}'$, und es sei die Hermitesche Form

$$H = \overline{\mathfrak{y}}' \mathfrak{A} \mathfrak{y} > 0$$

für beliebige Vektoren $\mathfrak{y}$ mit Ausnahme des Nullvektors. Bei einer solchen Matrix möge das Einzelschrittverfahren untersucht werden. Im Reellen ist die Konvergenz des Verfahrens bekannt; es soll hier gezeigt werden, daß die Konvergenzaussage sofort aus dem allgemeinen Satz in Nr. 1 folgt.

Es sei $\varkappa$ eine Wurzel der Gleichung

$$\det(\varkappa \mathfrak{B} + \mathfrak{C}) = 0. \tag{18}$$

Es gibt dann also einen vom Nullvektor verschiedenen Vektor $\mathfrak{x}$ mit

$$\varkappa \mathfrak{B} \mathfrak{x} + \mathfrak{C} \mathfrak{x} = 0. \tag{19}$$

Die Matrix $\mathfrak{B}$ läßt sich darstellen in der Form

$$\mathfrak{B} = \mathfrak{D} + \overline{\mathfrak{C}}',$$

wobei $\mathfrak{D}$ eine Diagonalmatrix ist, deren Diagonalelemente reell (weil $\mathfrak{A}$ hermitesch) und sogar positiv sind (weil $\mathfrak{A}$ positiv definit sein sollte). Schematisch wird also $\mathfrak{A}$ zerlegt:

$$\mathfrak{A} = \begin{pmatrix} & \mathfrak{C} \\ & \mathfrak{D} \\ \overline{\mathfrak{C}}' & \end{pmatrix}.$$

Mit $\varkappa = 1 - 2\nu$ wird somit nach (19)

$$\mathfrak{A} \mathfrak{x} = (\mathfrak{B} + \mathfrak{C}) \mathfrak{x} = (1 - \varkappa) \mathfrak{B} \mathfrak{x} = 2\nu \mathfrak{B} \mathfrak{x}$$

und

$$H = \overline{\mathfrak{x}}' \mathfrak{A} \mathfrak{x} = 2\nu \overline{\mathfrak{x}}' \mathfrak{B} \mathfrak{x} = \nu \overline{\mathfrak{x}}' \mathfrak{D} \mathfrak{x} + \nu \overline{\mathfrak{x}}' \mathfrak{A} \mathfrak{x} + \nu \overline{\mathfrak{x}}' (\overline{\mathfrak{C}}' - \mathfrak{C}) \mathfrak{x}.$$

$\mathfrak{D}$ ist eine positiv definite Matrix, also ist die Größe

$$D = \overline{\mathfrak{x}}' \mathfrak{D} \mathfrak{x}$$

positiv. Die Matrix $\mathfrak{G} = \overline{\mathfrak{C}}' - \mathfrak{C}$ ist schiefhermitesch, also

$$G = \overline{\mathfrak{x}}' (\overline{\mathfrak{C}}' - \mathfrak{C}) \mathfrak{x} = ig$$

rein imaginär. Somit wird

$$\nu = \frac{H}{D+H+ig}, \tag{20}$$

wobei $H>0$, $D>0$, und g reell ist. Lassen wir bei gegebenen Werten $H>0$, $D>0$ die Größe g alle reellen Zahlen durchlaufen, so beschreibt die zugehörige Größe ν in der Gaussschen Zahlenebene einen auf der reellen Achse in den Punkten O und $\frac{H}{D+H}$ senkrecht stehenden Kreis und $\varkappa = 1-2\nu$ einen die reelle Achse in den Punkten $\frac{D-H}{D+H}$ und 1 senkrecht schneidenden Kreis, vgl. Fig. 1; der zweitgenannte Kreis liegt wegen

$$-1 < \frac{D-H}{D+H} < 1$$

im Innern des Einheitskreises und hat mit diesem nur den Punkt 1 gemeinsam, der aber von $\varkappa$ nur für $g=\infty$ angenommen werden würde, was hier nicht in Betracht kommt. Es liegt also jede Wurzel $\varkappa$ von (18) innerhalb des Einheitskreises. Die Voraussetzungen des Satzes von Nr. 1 sind erfüllt, und es gilt

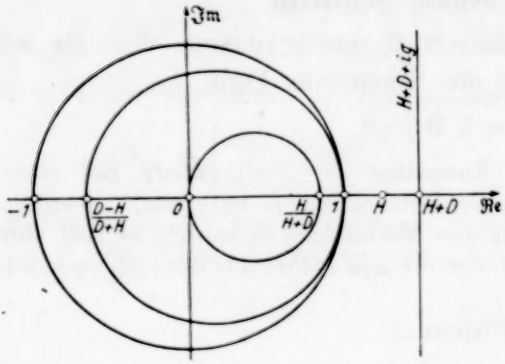

Fig. 1.

Ist im vorgelegten Gleichungssystem (1) *die Koeffizientenmatrix hermitesch und positiv definit, so konvergiert das Einzelschrittverfahren.*

Für eine Hermitesche Matrix $\mathfrak{A}$ ist zur Konvergenz des Einzelschrittverfahrens die positive Definitheit keineswegs notwendig, wie schon das einfache Beispiel

$$\mathfrak{A} = \begin{pmatrix} 1 & \varepsilon \\ \varepsilon & -1 \end{pmatrix}$$

mit $-1<\varepsilon<1$ zeigt.

Die drei Aussagen für eine Matrix

α) Das Einzelschrittverfahren konvergiert bei beliebigem Ausgangsvektor,

β) das Gesamtschrittverfahren konvergiert bei beliebigem Ausgangsvektor,

γ) $\mathfrak{A}$ ist hermitesch und positiv definit

überschneiden sich teilweise, wie das Beispiel der Matrix

$$\mathfrak{A} = \begin{pmatrix} 1 & a & b \\ |a| & 1 & a \\ |b| & |a| & 1 \end{pmatrix}$$

für reelle Werte von a, b zeigt. In einem rechtwinkligen a-b-Koordinatensystem kann man den Wertepaaren a, b, die zu Matrizen

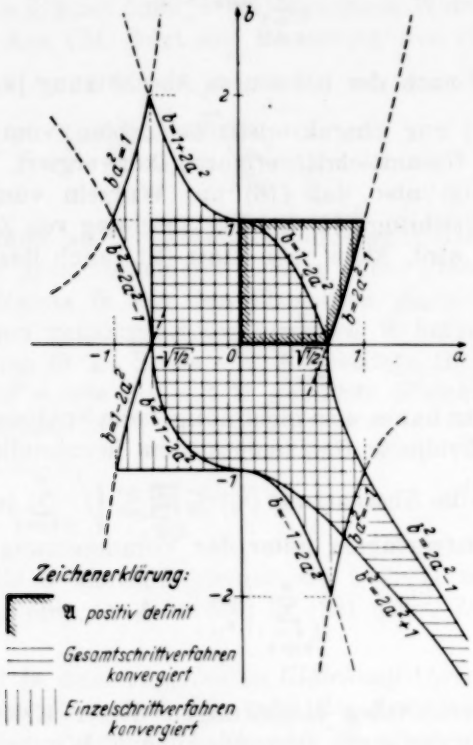

Fig. 2.

mit der Eigenschaft α) gehören, die entsprechenden Punkte der a-b-Ebene und somit der Eigenschaft α) ein Gebiet zuordnen, wobei nur die inneren Punkte gezählt werden, ebenso bei β) und γ). Wie es sein muß, gehört das Gebiet bei γ) ganz zu dem Gebiet bei α), aber sonst überschneiden sich die Gebiete; für jeden der sechs möglichen Fälle kann man ein Gebiet heraussuchen, vgl. Fig. 2.

3. Gewöhnliches Zeilensummen- und Spaltensummenkriterium.

In der Matrix $\mathfrak{A}$ sei in jeder Zeile der Betrag des Hauptdiagonalgliedes größer als die Summe der Beträge der übrigen $n-1$ Elemente

$$\sum_{\substack{k=1 \\ k \neq j}}^{n}{}' |a_{jk}| < |a_{jj}| \qquad (j = 1, \ldots, n). \tag{21}$$

Wir sagen dann, das Zeilensummenkriterium sei erfüllt. Der Strich an der Summe deutet (wie üblich) auf das Auslassen des Gliedes mit $k = j$ hin.

Im durchdividierten Gleichungssystem gilt dann für die Matrix $\mathfrak{A}^*$

$$\sum_{\substack{k=1\\k\neq j}}^{n}{}' |a_{jk}^*| < 1,$$

also hat $\mathfrak{A}^* - \mathfrak{E}$ nach der bekannten Abschätzung $|\varkappa_a| \leqq \max_j \sum_{k=1}^{n} |a_{jk}|$ (vgl. Fußnote [2]) nur charakteristische Zahlen vom Betrage kleiner als 1, und das Gesamtschrittverfahren konvergiert. Aus der Voraussetzung (21) folgt also, daß (16) nur Wurzeln vom Betrage kleiner als 1 hat. Da Gleichung (16) bei Vertauschung von Zeilen und Spalten nicht geändert wird, kann man statt (21) auch das Spaltensummenkriterium

$$\sum_{\substack{j=1\\j\neq k}}^{n}{}' |a_{jk}| < |a_{kk}| \qquad (k = 1, \ldots, n) \tag{22}$$

voraussetzen. Man hat es also nicht nötig, beim Spaltensummenkriterium erst das durchdividierte Gleichungssystem herzustellen.

Benutzt man die Abschätzung $|\varkappa_a| \leqq \overline{\mathfrak{A}} \leqq \sqrt{\sum_{j,k=1}^{n} |a_{jk}|^2}$, so folgt, daß das Gesamtschrittverfahren unter der Voraussetzung

$$\sum_{\substack{j,k=1\\j\neq k}}^{n}{}' \left|\frac{a_{jk}}{a_{jj}}\right|^2 < 1$$

konvergent ist.

Bei den gewöhnlichen Herleitungen dieser Kriterien pflegt man das Gesamtschrittverfahren zugrundezulegen. Wir betrachten nun das Einzelschrittverfahren und setzen zunächst (21) voraus.

Wir behaupten, aus (21) folgt, daß alle Wurzeln $\varkappa$ von (17) Beträge kleiner als 1 haben.

Beweis. Entweder sind alle Wurzeln $\varkappa$ von (17) gleich Null, dann ist die Behauptung richtig, oder es gibt eine Wurzel $\varkappa \neq 0$ von (17). Zu einer solchen gibt es einen vom Nullvektor verschiedenen Vektor $\mathfrak{x}$ mit den Komponenten x_j, so daß gilt:

$$\varkappa \sum_{k=1}^{\sigma} a_{\sigma k}\, x_k = -\sum_{k=\sigma+1}^{n} a_{\sigma k}\, x_k \qquad (\sigma = 1, 2, \ldots, n). \tag{23}$$

Unter den $|x_j|$ gibt es einen größten Betrag ξ, etwa $\xi = |x_s| > 0$, $\xi \geqq |x_j|$ für alle j. Gleichung (23) werde nun für $\sigma = s$ angeschrieben; es folgt

$$|\varkappa|\left(|a_{ss}|\,\xi - \sum_{k=1}^{s-1} |a_{sk}|\,\xi\right) \leqq |\varkappa| \left|\sum_{k=1}^{s} a_{sk}\, x_k\right| \tag{24}$$

$$= \left|\sum_{k=s+1}^{n} a_{sk}\, x_k\right| \leqq \sum_{k=s+1}^{n} |a_{sk}|\,\xi.$$

Mit der üblichen Verabredung, daß eine leere Summe durch die Zahl Null zu ersetzen ist, würde diese Ungleichung für den Fall $s = n$ wegen $|\varkappa| > 0, \xi > 0$ infolge (21) zu einem Widerspruch führen, es ist also $s < n$. Aus (24) folgt mit Benutzung von (21)

$$|\varkappa| \leqq \frac{\sum_{k=s+1}^{n} |a_{sk}|}{|a_{ss}| - \sum_{k=1}^{s-1} |a_{sk}|} < 1. \tag{25}$$

Durch Spiegelung an der Nebendiagonale geht das Zeilensummenkriterium in das Spaltensummenkriterium über. Präziser formuliert: Wir bilden die Matrix $\mathfrak{G}$ mit den Elementen $g_{jk} = a_{n+1-k,\, n+1-j}$, die durch die eben genannte Spiegelung aus $\mathfrak{A}$ hervorgeht. Die mit den Elementen von $\mathfrak{G}$ an Stelle von $\mathfrak{A}$ gebildete Gleichung (17) hat dieselben Wurzeln $\varkappa$ wie die mit $\mathfrak{A}$ gebildete Gleichung (17). Wenn $\mathfrak{G}$ also die Bedingung (21)

$$\sum_{\substack{k=1 \\ k \neq j}}^{n}{}' |g_{jk}| < |g_{jj}| \qquad (j = 1, 2, \ldots, n) \tag{26}$$

erfüllt, haben alle $\varkappa$ Beträge kleiner als 1. Setzt man in (26) die Definition der g_{jk} ein, so geht (26) in (22) über. Zusammenfassend gilt also der

Satz. *Erfüllt in dem vorgelegten Gleichunyssystem* (1) *die Matrix* $\mathfrak{A}$ *das Zeilensummenkriterium* (21) *oder das Spaltensummenkriterium* (22), *so konvergieren Gesamtschritt- und Einzelschrittverfahren.*

4. Schwaches Zeilensummen- und Spaltensummenkriterium.

Wenn bei einer Matrix

$$\sum_{\substack{k=1 \\ k \neq j}}^{n}{}' |a_{jk}| \quad \begin{cases} \leqq |a_{jj}| & \text{für } j = 1, 2, \ldots, n \\ < |a_{jj}| & \text{für mindestens ein } j = j_0 \end{cases} \tag{27}$$

gilt, so sagen wir, es sei das „schwache Zeilensummenkriterium" erfüllt. Dieses reicht nicht dafür aus, daß alle Wurzeln $\varkappa$ von (16) oder (17) Beträge kleiner als 1 haben, wie etwa das Beispiel

$$\mathfrak{A} = \begin{pmatrix} 1 & 0{,}1 & 0{,}1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

zeigt. Dazu muß man noch voraussetzen, daß die Matrix $\mathfrak{A}$ nicht „zerfällt". Wenn man bei einer Matrix $\mathfrak{A}$ die Indizes $1, \ldots, n$ so mit

$\varrho_1, \varrho_2, \ldots, \varrho_r, \sigma_1, \sigma_2, \ldots, \sigma_{n-r}$ (mit $1 \leqq r \leqq n-1$) numerieren kann, daß

$$a_{\varrho_\nu \sigma_\mu} = 0 \quad \text{für} \quad \nu = 1, 2, \ldots, r, \ \mu = 1, 2, \ldots, n-r \tag{28}$$

gilt, so heiße die Matrix „zerfallend".

Es sei nun $\mathfrak{A}$ nichtzerfallend und (27) erfüllt. Wir zeigen, daß dann alle Wurzeln $\varkappa$ von (16) und (17) Beträge kleiner als 1 haben, also Gesamtschritt- und Einzelschrittverfahren konvergieren. Es sei zunächst $\varkappa$ eine von Null verschiedene Wurzel von (16). Dann gibt es einen vom Nullvektor verschiedenen Vektor $\mathfrak{x}$ mit den Komponenten x_j und

$$|\varkappa|\,|a_{jj}|\,|x_j| = \left|\sum_{\substack{k=1\\k\neq j}}^{n}{}' a_{jk} x_k\right| \qquad (j = 1, \ldots, n). \tag{29}$$

Ist wie oben $\xi = |x_s| = \operatorname{Max}_j |x_j| > 0$, so folgt aus

$$\xi\,|\varkappa|\,|a_{ss}| \leqq \sum_{k}{}' |a_{sk}|\,\xi$$

nach (27) sofort $|\varkappa| \leqq 1$. Wir beweisen nun die Behauptung $|\varkappa| < 1$ indirekt. Es sei also $|\varkappa| = 1$. Aus (27) und (29) erhält man dann

$$\sum_{k}{}' |a_{sk}|\,|x_s| \leqq |a_{ss}|\,|x_s| \leqq \sum_{k}{}' |a_{sk}|\,|x_k| \tag{30}$$

oder

$$\sum_{k}{}' |a_{sk}|\,(|x_s| - |x_k|) \leqq 0\,. \tag{31}$$

Wegen $|x_s| \geqq |x_k|$ ist dies nur möglich, wenn $|x_s| = |x_k|$ ist für alle k, für die $a_{sk} \neq 0$ ist. Wären alle $a_{sk} = 0$ für $k \neq s$, so wäre die Matrix zerfallend. Es gibt also mindestens einen von s verschiedenen Index $k = \varrho_2$ mit $\xi = |x_{\varrho_1}| = |x_{\varrho_2}|$, wobei wir ϱ_1 für s geschrieben haben. Die gleiche Betrachtung für ϱ_2 ergibt, daß $\xi = |x_k|$ wird für alle k, für die $a_{\varrho_2 k} \neq 0$ ist. Bei Fortsetzung dieser Betrachtung sind zwei Fälle möglich: A) es sind alle $|x_k| = \xi$, B) es werden nur gewisse x_k erfaßt, etwa $|x_{\varrho_1}| = |x_{\varrho_2}| = \cdots = |x_{\varrho_r}| = \xi$, und in jeder für $s = \varrho_\nu$ angeschriebenen Gleichung tritt keines der übrigen x_k, die wir etwa $x_{\sigma_1}, x_{\sigma_2}, \ldots, x_{\sigma_{n-r}}$ nennen, auf, d. h. es ist

$$a_{\varrho_\nu \sigma_\mu} = 0 \quad \text{für alle} \quad \nu = 1, \ldots, r, \ \mu = 1, \ldots, n-r.$$

Der Fall B) ist aber auszuschließen, weil dann die Matrix zerfällt, und im Fall A) führt die Gleichung (29) für $j = j_0$ wegen $|\varkappa| = 1$, $|x_j| = |x_k| = \xi > 0$ unmittelbar zu einem Widerspruch mit (27).

Genau wie in Nr. 3 folgt, daß man die Rollen von Zeilen und Spalten vertauschen, also (27) durch das „schwache Spaltensummenkriterium"

$$\sum_{\substack{j=1\\j\neq k}}^{n}{}' |a_{jk}| \quad \begin{cases} \leqq |a_{kk}| & \text{für } k = 1, 2, \ldots, n \\ < |a_{kk}| & \text{für mindestens ein } k = k_0 \end{cases} \tag{32}$$

ersetzen kann.

Nun zeigen wir, daß unter denselben Voraussetzungen für die Matrix $\mathfrak{A}$ auch das Einzelschrittverfahren konvergiert, wobei wir uns etwas kürzer fassen. Es sei nun $\varkappa$ eine von Null verschiedene Wurzel von (17) und wieder $\mathfrak{x}$ ein Vektor mit nicht sämtlich verschwindenden Komponenten x_j, welche (23) erfüllen. Wieder sei $\xi = |x_s| = \operatorname{Max}_j |x_j|$; es folgt (24) und daraus hier zunächst nur $|\varkappa| \leqq 1$ an Stelle des früheren $|\varkappa| < 1$. Um zu zeigen, daß aber auch jetzt $|\varkappa| < 1$ gilt, wird die Annahme $|\varkappa| = 1$ zu einem Widerspruch geführt. Aus (23) ergibt sich (30), (31) und weiterhin wörtlich die gleiche Schlußweise wie oben beim Gesamtschrittverfahren.

Satz. *Ist in dem vorgelegten Gleichungssystem* (1) *die Matrix* $\mathfrak{A}$ *nicht zerfallend und außerdem das schwache Zeilensummenkriterium* (27) *oder das schwache Spaltensummenkriterium* (32) *erfüllt, so konvergieren Gesamtschritt- und Einzelschrittverfahren.*

5. Anwendungen des schwachen Zeilensummenkriteriums.

Das schwache Zeilensummenkriterium kann man bei Anwendungen sehr häufig benutzen. Wir greifen kurz einige Beispiele heraus.

1. Der Randwertaufgabe

$$-\frac{d}{dx}\left(f(x)\frac{dy}{dx}\right) + g(x)\,y = r(x), \qquad y(a) = y_a, \qquad y(b) = y_b$$

mit stetigem $g(x)$, $r(x)$, stetig differenzierbarem $f(x)$ und $f(x) > 0$, $g(x) \geqq 0$ in $a \leqq x \leqq b$ werden bei näherungsweiser Lösung mit dem gewöhnlichen Differenzenverfahren die Differenzengleichungen

$$(33)\qquad \frac{1}{h^2}\left[-f_{j+\frac{1}{2}} \cdot Y_{j+1} + \left(f_{j+\frac{1}{2}} + f_{j-\frac{1}{2}}\right) Y_j - f_{j-\frac{1}{2}} \cdot Y_{j-1}\right] + g_j Y_j = r_j \qquad (j = 1, \ldots, n-1),$$

$$Y_0 = y_a, \quad Y_n = y_b$$

zugeordnet. Dabei ist das Intervall a, b in n gleiche Teile der Länge $h = \frac{1}{n}(b-a)$ zerlegt, und der Wert einer Funktion $f, g, \ldots$ an der Stelle $x_k = a + k\,h$ wie üblich durch Anhängen des Index k gekennzeichnet, wobei k nicht ganzzahlig zu sein braucht. Y_j bezeichnet den durch Lösung des linearen Gleichungssystems (33) erhaltenen Näherungswert für die exakte Lösung $y(x_j)$. Diese Auflösung kann nach dem Satz in Nr. 4 iterativ in Gesamt- oder Einzelschritten erfolgen, da im Falle $g > 0$ das Zeilensummenkriterium in der gewöhnlichen Form und im Falle $g = 0$ in der schwachen Form erfüllt ist und die Matrix nicht zerfällt.

2. Bei der elliptischen Differentialgleichung

$$(34)\qquad L[u] = a\frac{\partial^2 u}{\partial x^2} + c\frac{\partial^2 u}{\partial y^2} + d\frac{\partial u}{\partial x} + e\frac{\partial u}{\partial y} - g\,u = r$$

seien die gegebenen Funktionen a, c, d, e, g, r in einem Bereich der x-y-Ebene stetig, a, c positiv und g nichtnegativ. In einem rechteckigen Gitter mit den Maschenweiten h, l, den Gitterpunkten

$$x_j = x_0 + jh \qquad y_k = y_0 + kl \qquad (j, k = 0, 1, 2, \ldots)$$

und Bezeichnung der Funktionswerte an einer Stelle x_j, y_k durch Anhängen der Indizes j, k entspricht der Differentialgleichung (34) die gewöhnliche Differenzengleichung für die Näherungswerte U_{jk}

$$\left(\frac{2a_{jk}}{h^2} + \frac{2c_{jk}}{l^2} + g_{jk}\right) U_{jk} - \frac{1}{h^2}\left(a_{jk} - \frac{h}{2} d_{jk}\right) U_{j-1,k} - \frac{1}{h^2}\left(a_{jk} + \frac{h}{2} d_{jk}\right) U_{j+1,k}$$
$$- \frac{1}{l^2}\left(c_{jk} - \frac{l}{2} e_{jk}\right) U_{j,k-1} - \frac{1}{l^2}\left(c_{jk} + \frac{l}{2} e_{jk}\right) U_{j,k+1} + r_{jk} = 0 .$$

Sind h und l so klein gewählt, daß alle runden Klammern nichtnegativ sind, so ist in jeder dieser Gleichungen der Koeffizient bei U_{jk} mindestens so groß wie die Summe der Beträge der Koeffizienten bei den übrigen Größen U. Die Differenzengleichungen werden angeschrieben für die inneren Punkte eines Gitterbereiches, wobei wir hier nicht auf die Definitionen von „Gitterbereich", „inneren Punkten" und „Randpunkten" eingehen, sondern auf die Literatur verweisen.

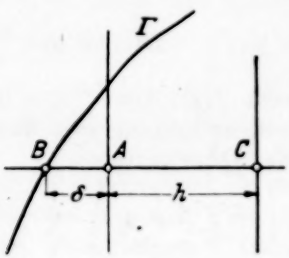

Fig. 3.

Sind längs einer Randkurve Γ die Werte von u vorgeschrieben, so treten mit den Bezeichnungen der Fig. 3 noch Gleichungen der Form hinzu

$$(\delta + h)\, u(A) - \delta\, u(C) = h\, u(B),$$

wobei $u(B)$ gegeben ist und der Koeffizient bei $u(A)$ größer ist als der Betrag des Koeffizienten $u(C)$ (oder, sofern $\delta = 0$ ist, in der Gleichung für den Punkt C der Wert $u(A)$ bekannt und damit ein Überwiegen des Betrages beim Koeffizienten von $u(C)$ gesichert ist), so daß das schwache Zeilensummenkriterium erfüllt ist, vorausgesetzt, daß der Bereich der inneren Gitterpunkte nicht in getrennte Bereiche und damit auch die Koeffizientenmatrix zerfällt.

3. Auch bei der 2. Randwertaufgabe für die elliptische Differentialgleichung (34) und bei partiellen Differentialgleichungen mit mehr als zwei unabhängigen Veränderlichen führt das Differenzenverfahren häufig auf Gleichungen, bei denen das schwache Zeilensummenkriterium unmittelbar die Konvergenz von Gesamtschritt- und Einzelschrittverfahren sichert.

In diesen genannten Fällen sind viele in der Literatur behandelte Fälle[6]) als Spezialfälle enthalten, besonders genannt seien nur das Liebmannsche Mittelungsverfahren und die Relaxation von Southwell.

[6]) H. Liebmann, Die angenäherte Ermittlung harmonischer Funktionen und konformer Abbildung, Sitz. Ber. math. phys. Klasse, bayr. Akad. Wiss. München 1918, S. 385—416.

R. Courant, Über Randwertaufgaben bei partiellen Differentialgleichungen, Z. angew. Math. Mech. 6 (1926), 322—325.

R. V. Southwell, Relaxation methods in theoretical physics, Oxford 1946.

(Eingegangen am 12. Februar 1950.)

Mathematische Zeitschrift, Band 53, Heft 2, S. 162—201 (1950).

Konstruktive Begründung der Mathematik.

Herrn Erich Kamke zum 60. Geburtstag am 18. August 1950 gewidmet.

Von

Paul Lorenzen in Bonn.

Das Problem der Widerspruchsfreiheit der Mathematik hat nur dann einen Sinn, wenn es auf einen vorgegebenen Kalkül bezogen ist. Ist K ein Kalkül, A die Menge seiner Aussagen und S die Menge seiner Sätze, so heißt K widerspruchsfrei, wenn A eine echte Untermenge von S ist, wenn also nicht alle Aussagen Sätze sind.

Nach dem Hilbertschen Programm zur Grundlegung der Mathematik wird die Aufgabe der Widerspruchsfreiheitsbeweise einer Metamathematik zugewiesen. Diese Auffassung des Problems hat in der Tat zu Beweisen geführt, die mit dem sogenannten inhaltlichen Schließen zu der Aussage der Widerspruchsfreiheit, z. B. für die Arithmetik führen. Da nun jeder Beweis nur so viel taugt, wie die Mittel, mit denen er geführt wird, bleiben diese Beweise dem Einwand ausgesetzt, daß das inhaltliche Schließen weder formalisiert, geschweige denn begründet ist.

Bei dem Versuch einer solchen Formalisierung hat sich herausgestellt, daß die Unterscheidung zwischen Mathematik und Metamathematik dem Grundlagenproblem nicht angemessen ist. Die Beweismittel der Metamathematik sind keine anderen als die der Mathematik. Eine Begründung der Metamathematik ist daher nichts anderes als die Begründung eines Teiles der Mathematik, (etwa dessen, was man intuitionistische Mathematik nennt). Es ist weiter bekannt, daß z. B. aus der Widerspruchsfreiheit der intuitionistischen Logik die Widerspruchsfreiheit der zweiwertigen, der sog. klassischen Logik trivial zu beweisen ist; d. h. wenn die intuitionistische Arithmetik begründet wäre, so wäre auch die klassische leicht als eine brauchbare Erweiterung zu begründen.

Ein einwandfreier, und zugleich gegenüber den bisherigen, wesentlich einfacherer Weg zur Widerspruchsfreiheit der Arithmetik ergibt sich also, wenn die intuitionistische Mathematik in formalisierter Gestalt begründet werden kann. Entgegen dem mißverständlichen Namen „intuitionistisch" läßt sich diese Grundlegung nun in der Tat ohne irgend welche Intuitionen, die die fehlenden Beweise ersetzen müßten, durchführen.

I. Wir gehen aus von Regeln zur Konstruktion von Zeichen, z. B.

(Z_1) $\quad |$

(Z_2) $\quad x \underset{x}{\rightarrow} x|$

zur Konstruktion der Zeichen

$$|, ||, |||, \ldots$$

(umgangssprachlich könnte etwa $\rightarrow$ als „dann", die darunter befindliche Variable als „für jedes x" gelesen werden).

Oder z. B.

(D_1) $\quad | = |$

(D_2) $\quad z_1 = z_2 \underset{z_1, z_2}{\rightarrow} z_1| = z_2|$

zur Konstruktion der Zeichen

$$| = |, || = ||, ||| = |||, \ldots.$$

Solche Regeln werden nicht darauf befragt, ob sie „wahr" sind oder nicht — sondern nur ob sie „brauchbar" sind oder nicht, d. h. ob das Handeln nach ihnen, also das Konstruieren der Zeichen für irgend einen Zweck brauchbar ist. Das ist aber keine mathematische Frage mehr, sondern gehört zu den Anwendungen. Das Verstehen dieser Regeln ist durch einfaches Vor- und Nachmachen des Konstruierens nach ihnen zu lernen, es ist unabhängig vom Verstehen einer der traditionellen Umgangssprachen. In diesem Sinne ist die hier gegebene Begründung „sprachunabhängig".

Zur Erleichterung des Ableitens, d. h. des Konstruierens von Zeichen nach gegebenen Regeln werden weitere Regeln aufgestellt, deren Benutzung eliminiert werden kann: sind z. B. zwei Regeln $p_1 \rightarrow p_2$ und $p_2 \rightarrow p_3$ gegeben, dann kann die Benutzung der Regel $p_1 \rightarrow p_3$ aus jeder Ableitung durch die Benutzung der gegebenen Regeln eliminiert werden. Durch Angabe des Eliminationsverfahrens in einer geeigneten Symbolik, ist damit von der Metaregel [1])

$$p_1 \rightarrow p_2; \quad p_2 \rightarrow p_3 \therefore p_1 \rightarrow p_3$$

bewiesen, daß sie die Menge der gegebenen Regeln nur um eliminierbare Regeln erweitert — grob gesagt: daß sie nichts Neues liefert. Durch ähnliche Eliminationsverfahren werden alle Sätze der sog. Konsequenzlogik beweisbar. Die Beweise sind zwar keine formalen Ableitungen, wohl aber formale Operationen, die stets mechanisch kontrolliert werden können — in diesem Sinne ist die Grundlegung allgemein gültig, von jeder subjektiven Meinung unabhängig.

II. Zur Vereinfachung der konsequenzlogischen Sätze empfiehlt sich dann die Einführung formaler Aussageverknüpfungen, insbesondere

[1]) Wir benutzen Punkte über, und $\rightarrow$ statt Klammern.

von Konjunktion $\frown$ und Disjunktion $\smile$. Auch hier ist kein Bezug auf die Umgangssprache erforderlich (die Zeichen brauchen nicht übersetzt zu werden) — und die „Einsicht" wird nur für die kalkulatorische Zweckmäßigkeit, die Anwendbarkeit, nicht für mathematische Aussagen gebraucht. Statt einer Negation wird zunächst nur „das Falsche" eingeführt, d. h. ein Zeichen $\wedge$ das durch jede Aussage ersetzt werden darf, die nach den gegebenen Regeln unableitbar ist. Dabei ist unableitbar definiert als: „ungleich allen ableitbaren Aussagen", wird also zurückgeführt auf die Ungleichheit von Zeichen, auf eine formal-entscheidbare Aussage. Jetzt kann die Negation $\bar{p}$ einer Aussage p definiert werden durch $p \to \wedge$, und die Sätze der intuitionistischen Logik wie

$$p_1 \to p_2 \dashrightarrow \bar{p}_2 \to \bar{p}_1,$$
$$p \to \bar{\bar{p}}$$

werden aus der Konsequenzlogik ableitbar. Man erhält dagegen nicht das sog. tertium non datur: $p \smile \bar{p}$.

Bildet man jedoch jede Aussage p auf eine geeignete Aussage p' ab, indem man die disjunktionsfreien Aussagen auf sich selber abbildet, aber z. B. $p_1 \smile p_2$ auf

$$\overline{\overline{p_1' \smile p_2'}},$$

so wird wenigstens

$$(p \smile \bar{p})'$$

ableitbar. Die Hinzunahme des tertium non datur, d. h. der Übergang von der intuitionistischen zur zweiwertigen Logik ändert hiernach an der Ableitbarkeit der disjunktionsfreien Aussagen gar nichts, insbesondere auch nichts an der Unableitbarkeit von $\wedge$.

III. Das Problem der Widerspruchsfreiheit der Arithmetik wird fälschlich oft als Frage nach der Widerspruchsfreiheit eines Axiomensystems, etwa des Peano-Systems, gestellt. Wenn die Zahlen z als die nach (Z_1), (Z_2) konstruierbaren Zeichen definiert sind, und die Gleichheiten $z_1 = z_2$ als die nach (D_1), (D_2) konstruierbaren Zeichen, so ist für das „Axiom"

$$z_1| = z_2| \to z_1 = z_2$$

zu beweisen, daß es eine eliminierbare Regel ist. Für das Axiom

$$\overline{z| = |}$$

ist zu beweisen, daß $z| = |$ unableitbar ist, und die Induktionsregel

$$p(|); \quad p(z) \underset{z}{\dashrightarrow} p(z|) \dashrightarrow p(z)$$

ist als Metaregel zu beweisen, d. h. es ist ein Verfahren anzugeben, nach dem $p(z)$ eliminierbar ist, wenn $p(|)$ und $p(z) \underset{z}{\dashrightarrow} p(z|)$ eliminierbar sind.

Daß diese Beweise trivial sind, zeigt, daß die Frage nach der Widerspruchsfreiheit der Axiome falsch gestellt ist. (Die oft stillschweigend eingeführten Axiome

$$z = z,$$
$$z_1 = z_2, \quad p(z_1) \to p(z_2)$$

lassen sich sogar ableiten.)

Nachdem die Widerspruchsfreiheit der Logik (problematisch ist hier nur das tertium non datur) gezeigt ist, bleibt für die arithmetischen Axiome nur die Angabe eines konstruktiven Modelles, d. h. eines Systems von Regeln zur Konstruktion von Zeichen, so daß die gegebenen Axiome beweisbar werden.

IV. Während dieses konstruktive Modell für die Arithmetik trivial ist, liegt für die Kalküle der axiomatischen Analysis und Mengenlehre gerade hier die Schwierigkeit. Ein Blick auf die Geschichte dieser Kalküle zeigt jedoch, daß diese als Axiomatisierungen von schon vorhandenen Konstruktionen entstanden sind. Die Axiomatisierungen haben keinen anderen Zweck, als das Operieren mit konstruierbaren Mengen, Funktionen, reellen Zahlen usw. zu erleichtern — genau so wie auch die axiomatischen Untersuchungen der Algebra höchstens dann sinnvoll sind, wenn es Modelle für die axiomatisch definierten Strukturen gibt. Zieht man in Betracht, daß die bisherigen Axiomatisierungsversuche immer stark beeinflußt waren von den umgangssprachlichen Denkgewohnheiten, so wird die Entstehung dieser Kalküle psychologisch verständlich. (Auch die traditionellen Sprachen kennen keinen konstruktiven Aufbau der Begriffe. Dies macht verständlich, daß man bei der Bildung der „Menge aller x mit der Eigenschaft $p(x)$" vergißt anzugeben, wie das $p(x)$ aussehen soll — und daher z. B. „Menge" unbemerkt in doppelter Bedeutung gebraucht, wenn in $p(x)$ etwa der Redeteil „für alle Mengen" vorkommt. Nur durch dieses Verwischen des schrittweisen Aufbaues von Mengen kommt die sog. Überabzählbarkeit zustande — da man andererseits trotzdem stets nur abzählbare Mengen vor sich hat, wird als Ersatz für die Abzählbarkeit dann das Auswahlaxiom verständlich.)

Obwohl es also gerechtfertigt wäre, axiomatische Mengenlehren, für die kein konstruktives Modell vorhanden ist, einfach als mißglückte Versuche in Zukunft beiseite zu lassen, besteht aber auch die andere Möglichkeit, zu untersuchen, ob sich solche Axiomatisierungen — nachdem sie einmal da sind und wir uns daran gewöhnt haben — nicht doch auf irgend einem Umweg als kalkulatorisch zweckmäßige Fiktionen für die konstruktive Mathematik verwenden lassen. Mit der bloßen Widerspruchsfreiheit der axiomatischen Mengenlehre wäre allerdings kaum etwas gewonnen, es kommt darauf an, den Zusammenhang mit der konstruktiven Mengenlehre herzustellen. Als naheliegendste Richtung zur Weiterarbeit ergibt sich hieraus: zunächst das, was hier

unter dem Namen „konstruktive Mengenlehre" zusammengefaßt wurde, ausführlich zu entwickeln. Im Gegensatz zu den intuitionistischen Versuchen darf dabei jetzt — nach einwandfreier Begründung der Logik — das tertium non datur stets benutzt werden. Die so unbequeme Beschränkung auf „entscheidbare Eigenschaften", „berechenbare reelle Zahlen" usw. ist nicht mehr erforderlich. Einen Ansatz in der hier gemeinten Richtung hat H. Weyl schon 1918 gemacht — dieser mußte allerdings scheitern, da die Begründung des tertium non datur damals noch fehlte.

Der Ausbau einer konstruktiven Analysis ist hierbei selbstverständlich auch unabhängig von dem Ziel, die Brauchbarkeit der traditionellen Axiomatisierungen zu beweisen, von Interesse. Es eröffnen sich neue Möglichkeiten, etwa für das formale Rechnen mit unendlichen Prozessen, unabhängig vom Begriff der reellen Zahl und der Konvergenz.

Durch die Ausnutzung solcher Möglichkeiten wie der transfiniten Iteration für den Prozeß der Mengenbildung ergeben sich vielleicht auch genügend bequeme Kalküle, daß die gegenwärtigen Axiomatisierungen einfach überflüssig werden.

Die vorliegende Arbeit beschränkt sich jedoch darauf, den Aufbau der Mathematik nur bis zur Arithmetik lückenlos durchzuführen.

§ 1. Die konsequenzlogische Symbolik.
§ 2. Konsequenzlogische Sätze.
§ 3. Definitionen und Inversionen. Das Induktionsprinzip.
§ 4. Positive Logik.
§ 5. Arithmetik.
§ 6. Semiotik.
§ 7. Unableitbarkeit.
§ 8. Negation.
§ 9. Tertium non datur.

§ 1.

Die konsequenzlogische Symbolik.

1. Wir verwenden $x, x_1, \ldots$ als Variable für Zeichen (endliche Reihen von Einzelzeichen).

Wir definieren Zahlen als die nach

(Z_1) $\quad |$

(Z_2) $\quad x \underset{x}{\rightarrow} x|$

konstruierbaren Zeichen.

Wir verwenden $z, z_1, z_2, \ldots$ als Variable für Zahlen.
Unter Aussagen verstehen wir die Zeichen

$$z_1 = z_2, \quad z_1 \neq z_2, \quad z_1 + z_2 = z_3, \quad z_1 \cdot z_2 = z_3,$$

d. h. diejenigen, die aus den eben angeschriebenen bei Ersetzung der Variablen durch Zahlen entstehen.

Als Sätze definieren wir die nach

(D_1) $\quad | = |$

(D_2) $\quad z_1 = z_2 \underset{z_1, z_2}{\rightarrow} z_1 | = z_2 |$

(D_3) $\quad z | \neq_z |$

(D_4) $\quad | \neq_z z |$

(D_5) $\quad z_1 \neq z_2 \underset{z_1, z_2}{\rightarrow} z_1 | \neq z_2 |$

(D_6) $\quad z_1 | = z_2 \underset{z_1, z_2}{\rightarrow} z_1 + | = z_2$

(D_7) $\quad z_1 + z_2 = z, \quad z | = z_3 \underset{z, z_1, z_2, z_3}{\rightarrow} z_1 + z_2 | = z_3$

(D_8) $\quad z_1 = z_2 \underset{z_1, z_2}{\rightarrow} z_1 \cdot | = z_2$

(D_9) $\quad z_1 \cdot z_2 = z, \quad z + z_1 = z_3 \underset{z, z_1, z_2, z_3}{\rightarrow} z_1 \cdot z_2 | = z_3$

konstruierbaren Zeichen. (Alle Sätze sind Aussagen.)

Der Zweck dieser Konstruktionen liegt in den Anwendungen und interessiert hier nicht.

Wir verwenden $p, p_1, p_2, \ldots$ als Variable für Aussagen. Später werden wir den Variabilitätsbereich von p erweitern, zunächst kommt es uns jedoch nur darauf an, daß wir ein System von Regeln zur Konstruktion gewisser Aussagen haben. Alle Regeln haben die Form

$$p_1, \ldots, p_n \rightarrow p,$$

wobei — wie in (D_1), (D_3), (D_4) — auch die Prämissen $p_1, \ldots, p_n$ fehlen können. (Dann lassen wir auch den Pfeil $\rightarrow$ weg.) Unter dem Pfeil — oder der alleinstehenden Konklusion wie in (D_3), (D_4) — sind evtl. noch gebundene Variable hinzuzufügen.

2. Das Konstruieren von Zeichen nach solchen Regeln nennen wir Ableiten. Das Ableiten kann durch hinreichend häufige Übung erlernt werden — pädagogische Einzelheiten interessieren hier jedoch ebenfalls nicht. Wir begnügen uns mit einigen Beispielen, durch die wir gleichzeitig für das Ableiten eine einfache Symbolik zur Verweisung auf die bei jedem Ableitungsschritt benutzten Regeln und Prämissen einführen.

1. Beispiel:

$$D_1 \vdash \quad 1) \quad \mathsf{I} = \mathsf{I}$$
$$D_2 \vdash 1 \to 2) \quad \mathsf{II} = \mathsf{II}$$
$$D_2 \vdash 2 \to 3) \quad \mathsf{III} = \mathsf{III}$$
$$D_2 \vdash 3 \to 4) \quad \mathsf{IIII} = \mathsf{IIII},$$

Ergebnis:

$$\vdash \mathsf{IIII} = \mathsf{IIII}.$$

2. Beispiel: (Fortsetzung von 1.)

$$D_4 \vdash 3 \quad \to 5) \quad \mathsf{II} + \mathsf{I} = \mathsf{III}$$
$$D_7 \vdash 5, 4 \to 6) \quad \mathsf{II} + \mathsf{II} = \mathsf{IIII}$$
$$D_8 \vdash 2 \quad \to 7) \quad \mathsf{II} \cdot \mathsf{I} = \mathsf{II}$$
$$D_9 \vdash 7, 6 \to 8) \quad \mathsf{II} \cdot \mathsf{II} = \mathsf{IIII},$$

Ergebnis:

$$\vdash \mathsf{II} \cdot \mathsf{II} = \mathsf{IIII}.$$

Links von $\vdash$ (gelesen etwa: also) steht die benutzte Regel. Rechts von $\vdash$ stehen die als Prämissen benutzten Aussagen (diese können evtl. fehlen).

Rechts von $\to$ folgt dann die Konklusion. Die Konklusionen sind durchlaufend numeriert, so daß alle Prämissen stets kleinere Nummern tragen müssen. (Die Nummern könnten auch durch geeignete Verbindungsstriche ersetzt werden.) Ersetzt man die Nummern durch die zugehörigen Aussagen, so steht rechts von $\vdash$ stets eine Spezialisierung der Regel links von $\vdash$. Z. B. würde die letzte Zeile im 2. Beispiel lauten:

$$z_1 \cdot z_2 = z, \quad z + z_1 = z_3 \underset{z,\, z_1,\, z_2,\, z_3}{\to} z_1 \cdot z_2 \mathsf{I} = z_3 \vdash \mathsf{II} \cdot \mathsf{I} = \mathsf{II},$$
$$\mathsf{II} + \mathsf{II} = \mathsf{IIII} \to \mathsf{II} \cdot \mathsf{II} = \mathsf{IIII}.$$

Eine Spezialisierung ensteht dabei durch Weglassen der Variablen unter der Regel und durch Ersetzung der restlichen Variablen durch Konstanten aus dem Variabilitätsbereich.

3. Wir erweitern die Herstellung von Ableitungen, indem wir die Benutzung von freien Variablen zulassen (bisher hatten wir nur gebundene Variable).

3. Beispiel:

$$D_1 \vdash \quad 1) \quad z\mathsf{I} \neq \mathsf{I}$$
$$D_3 \vdash 1 \to 2) \quad z\mathsf{II} \neq \mathsf{II}$$
$$D_3 \vdash 2 \to 3) \quad z\mathsf{III} \neq \mathsf{III},$$

Ergebnis:

$$\vdash z\mathsf{III} \neq \mathsf{III}.$$

Die Erweiterung kommt dadurch zustande, daß als Spezialisierung jetzt auch die Ersetzung der Variablen durch Variable, nicht nur durch Konstanten zugelassen ist. Zeichen mit freien Variablen, aus denen bei Ersetzung der Variablen durch Konstanten Aussagen entstehen, nennen wir Prädikate. Wir benutzen z. B. $p(z)$ als Variable für Prädikate, in denen z vorkommt. Durch Ersetzung entstehen hieraus die Aussagen $p(\mathrm{I})$, $p(\mathrm{II})$, $p(\mathrm{III})$, ...

Ersetzt man in einer Ableitung von $p(z)$ die Variable z überall durch eine Konstante, z. B. III, so entsteht eine Ableitung von $p(\mathrm{III})$. Eine Ableitung mit dem Ergebnis $p(z)$ benutzen wir daher als Beweis von $p(z)$. Die Einführung von freien Variablen, also von Prädikaten, ist der erste Kunstgriff, um Ableitungen zu ersparen. Um die Frage, ob etwa die Aussagen $p(\mathrm{III})$, $p(\mathrm{IIII})$ ableitbar sind, zu entscheiden, würde es genügen, $p(z)$ zu beweisen.

4. Insbesondere für die Ableitungen von Prädikaten ist es zweckmäßig, noch unvollständige Ableitungen einzuführen, indem man bei einzelnen Ableitungsschritten nicht nur schon abgeleitete Prämissen benutzt.

4. Beispiel:

$$D_8 \vdash \qquad z_1 = z_2 \to 1)\ z_1 \cdot \mathrm{I} = z_2$$
$$D_9 \vdash 1,\ z_2 + z_1 = z_3 \to 2)\ z_1 \cdot \mathrm{II} = z_3 .$$

Als Ergebnis dieser unvollständigen Ableitung formulieren wir

$$\vdash z_1 = z_2,\quad z_2 + z_1 = z_3 \to z_1 \cdot \mathrm{II} = z_3 .$$

Das Ergebnis einer unvollständigen Ableitung enthält rechts von $\vdash$ zunächst alle zusätzlich eingeführten Prämissen (in irgend einer Reihenfolge, aber gleiche nur einmal), dann folgt $\to$ und dann das abgeleitete Prädikat (evtl. Aussage). Aus dem Ergebnis erhalten wir eine Regel durch Weglassen von $\vdash$ und Bindung aller freien Variablen, z. B.

$$z_1 = z_2,\quad z_2 + z_1 = z_3 \underset{z_1,\ z_2,\ z_3}{\to} z_1 \cdot \mathrm{II} = z_3 .$$

Von den so erhaltenen Regeln kann folgender Gebrauch gemacht werden, um die Aufgabe, die Ableitbarkeit einer Aussage nach gegebenen Regeln zu untersuchen, zu erleichtern. Eine Ableitung, die außer den gegebenen Regeln noch die obige Regel benutzt, kann leicht in eine Ableitung nach den gegebenen Regeln allein umgeformt werden, d. h. die obige Regel ist eliminierbar. Enthält diese Ableitung nämlich die Zeile

$$z_1 = z_2,\quad z_2 + z_1 = z_3 \underset{z_1,\ z_2,\ z_3}{\to} z_1 \cdot \mathrm{II} = z_3 \vdash n_1,\ n_2 \to n)\ z_1 \cdot \mathrm{II} = z_3 ,$$

so enthält sie als schon abgeleitet oberhalb Zeilen mit

$$(n_1) \qquad z_1 = z_2$$

und ebenso

(n_2) $$z_2 + z_1 = z_3.$$

Ersetzt man die obige Zeile durch die unvollständige Ableitung, so entsteht wieder eine Ableitung.

Die Frage nach der Ableitbarkeit wird also dadurch erleichtert, daß man das System der gegebenen Regeln um eliminierbare Regeln erweitert.

Die unvollständigen Ableitungen liefern eliminierbare Regeln. Das Eliminationsverfahren wird durch die unvollständige Ableitung hinreichend deutlich beschrieben. Wir benutzen daher die unvollständige Ableitung kurz als Beweis der daraus zu erhaltenden eliminierbaren Regeln.

5. Beispiel:

$$z_1 = z_2 \underset{z_1,\, z_2}{\rightarrow} z_1\mathrm{III} = z_2\mathrm{III}.$$

Beweis.

$$\begin{aligned} D_2 &\vdash z_1 = z_2 \rightarrow 1) \quad z_1\mathrm{I} = z_2\mathrm{I} \\ D_3 &\vdash \quad 1 \quad \rightarrow 2) \quad z_1\mathrm{II} = z_2\mathrm{II} \\ D_3 &\vdash \quad 2 \quad \rightarrow 3) \quad z_1\mathrm{III} = z_2\mathrm{III}. \end{aligned}$$

5. Eine weitere Art von unvollständigen Ableitungen erhalten wir durch Benutzung von bisher unbewiesenen Regeln.

6. Beispiel:

$$\begin{aligned} z = z \underset{z}{\vdash} \quad & 1) \quad z\mathrm{I} = z\mathrm{I} \\ D_6 \vdash\ & 1 \rightarrow 2) \quad z+\mathrm{I} = z\mathrm{I}. \end{aligned}$$

Ergebnis:

$$z = z \underset{z}{\vdash} z + \mathrm{I} = z\mathrm{I}.$$

7. Beispiel: (Fortsetzung von 6.)

$$z_1 + z_2 = z_3 \underset{z_1,\, z_2,\, z_3}{\rightarrow} z_2 + z_1 = z_3 \vdash 2 \rightarrow 3)\ \mathrm{I} + z = z\mathrm{I},$$

Ergebnis:

$$z = z; \quad z_1 + z_2 = z_3 \underset{z_1, z_2, z_3}{\rightarrow} z_2 + z_1 = z_3 \vdash \mathrm{I} + z = z\mathrm{I}.$$

Das Ergebnis der unvollständigen Ableitung wird erhalten, indem links von $\vdash$ die unbewiesenen Regeln (in irgendeiner Reihenfolge, aber gleiche nur einmal) geschrieben werden (zur Abtrennung setzen wir ; zwischen die Regeln), rechts von $\vdash$ setzen wir die unabgeleiteten Prämissen, dann folgt $\rightarrow$ und dann die Konklusion. Mit Hilfe solcher unvollständigen Ableitungen können wir weitere eliminierbare Regeln erhalten, da sie ein Verfahren liefern, die Regeln rechts von $\vdash$ zu eliminieren, falls die Regeln links von $\vdash$ schon eliminiert werden können.

Es empfiehlt sich also, aus einer unvollständigen Ableitung eine Metaregel aufzustellen zur Konstruktion eliminierbarer Regeln, indem – durch $\dashrightarrow$ ersetzt wird und die freien Variablen gebunden werden. Wir benutzen auch hier die unvollständige Ableitung als Beweis der Metaregel.

8. Beispiel:

$$z_1 \cdot z_2 = z \underset{z_1, z_2}{\rightarrow} z_2 \cdot z_1 = z \underset{z_0}{\dashrightarrow} z_0 = z \rightarrow \mathsf{I} \cdot z_0 = z.$$

Beweis.

$$D_8 \vdash z_0 = z \rightarrow 1)\ z_0 \cdot \mathsf{I} = z$$

$$z_1 \cdot z_2 = z \underset{z_1, z_2}{\rightarrow} z_2 \cdot z_1 = z \vdash \quad 1 \quad \rightarrow 2)\ \mathsf{I} \cdot z_0 = z.$$

(Hier steht links ein Regelprädikat.) Der Beweis liefert ein Verfahren, wie das Regelprädikat

$$z_0 = z \underset{z_0}{\rightarrow} \mathsf{I} \cdot z_0 = z$$

aus einer Ableitung zu eliminieren ist durch Benutzung des Regelprädikates

$$z_1 \cdot z_2 = z \underset{z_1, z_2}{\rightarrow} z_2 \cdot z_1 = z.$$

6. Wir wollen von jetzt ab unter Aussage nicht nur die bisherigen Aussagen verstehen, sondern auch Prädikate, Regeln, Metaregeln, Metametaregeln, usw.

Wir benutzen die Variablen $p, p_1, p_2, \ldots$ auch für die Aussagen im weiteren Sinne. Die bisherigen Aussagen nennen wir Aussagen erster Stufe.

Unter Prädikat verstehen wir — wie bisher — ein Zeichen, das durch Ersetzung der freien Variablen in eine Aussage (jetzt: beliebiger Stufe) übergeht.

Die Regeln zur Konstruktion von Aussagen sind also folgende:

(P_1) $z_1 = z_2$, $z_1 \neq z_2$, $z_1 + z_2 = z_3$, $z_1 \cdot z_2 = z_3$ und alle Spezialisierungen.

(P_2) Gehört $p(z_1, \ldots, z_n)$ dazu, dann auch $p\underset{z_1, \ldots, z_n}{(z_1, \ldots, z_n)}$.

(P_3) Gehören $p, p_1, \ldots, p_n$ dazu und ist p nicht von der Form $p\underset{z_1, \ldots, z_n}{(z_1, \ldots, z_n)}$, dann gehört auch $p_1, \ldots, p_n \rightarrow p$ dazu.

(P_4) Gehört p dazu, dann ist p Prädikat oder Aussage, je nachdem ob p freie Variable enthält oder nicht.

(Es wäre leicht, mit Hilfe der später zu entwickelnden Semiotik diese Regeln (P_1) bis (P_4) zu formalisieren.)

Durch die Erweiterung des Begriffes Aussage kann es erforderlich werden, z. B. nach Regeln der Form $p\underset{z}{(z)} \rightarrow p_0$ abzuleiten.

9. Beispiel:

$$D_4 \vdash \quad 1)\ \mathsf{I} \neq z\,\mathsf{I}$$
$$D_5 \vdash 1 \to 2)\ \mathsf{II} \neq z\,\mathsf{II}$$
$$z_0\,\mathsf{I} \underset{z}{\neq} z\,\mathsf{II} \underset{z_0}{\to} z_0 = \mathsf{I} \vdash 2 \to 3)\ \mathsf{I} = \mathsf{I}.$$

Hier ist 2) das Prädikat $\mathsf{II} \neq z\,\mathsf{II}$. Um daraus die Prämisse der Spezialisierung von $z_0\,\mathsf{I} \underset{z}{\neq} z\,\mathsf{II} \underset{z_0}{\to} z_0 = \mathsf{I}$ (Ersetzung von z_0 durch I) zu erhalten, muß $\mathsf{II} \underset{z}{\neq} z\,\mathsf{II}$ gebildet werden. Die Ableitung ist also nur dann richtig, wenn die darin enthaltene Ableitung des Prädikates $\mathsf{II} \neq z\,\mathsf{II}$ tatsächlich $\mathsf{II} \underset{z}{\neq} z\,\mathsf{II}$ beweist. Dazu braucht man nicht etwa „alle" Ersetzungen von z durch $\mathsf{I}, \mathsf{II}, \mathsf{III}, \ldots$ durchzuprüfen (das wäre sinnlos), man braucht nur zu prüfen, daß die Ableitung des Prädikates $p(z)$ nicht davon abhängt, daß hier gerade z auftritt. Man ersetze z durch eine Variable, die sonst nirgends in der Ableitung vorkommt: bleibt die Ableitung richtig, dann beweist sie $p(z)$.

§ 2.

Konsequenzlogische Sätze.

Nachdem wir in § 1 an Beispielen aus der Arithmetik Methoden zum Beweis von Aussagen kennengelernt haben, wollen wir jetzt untersuchen, welche Aussagen sich beweisen lassen, wenn ein beliebiges System von Regeln vorgegeben ist. Wir wollen also nur voraussetzen, daß der Begriff „Aussage 1. Stufe" irgendwie definiert sei, und irgendwelche Regeln $p_1, \ldots, p_n \to p$ gegeben seien. Es mögen auch Variable $x, x_1, x_2, \ldots$ so eingeführt sein, daß die Aussagen durch Ersetzung der Variablen aus Prädikaten $p(x), p(x_1, x_2), \ldots$ entstehen.

Alle beweisbaren Aussagen nennen wir Sätze. Diejenigen Sätze, zu deren Beweis keinerlei Kenntnis der speziellen Regeln erforderlich ist, nennen wir konsequenzlogische Sätze; wir werden diese mit $L_1, L_2, \ldots$ bezeichnen.

(L_1) $\qquad p \to p$

Beweis. $p \vdash 1)\ p$

(L_2) $\qquad p_1, p_2 \to p_1$

Beweis. $p_1 \vdash 1)\ p_2$
$p_1 \vdash 2)\ p_1$

(L_3) $\qquad p_1, p_2 \to p_2$

Beweis. Entsprechend zu (L_2).

(L_4) $\qquad p(x) \to p(x)$

Beweis. $p(x) \underset{x}{\vdash} 1)\ p(x)$

(L_5) $\quad p_1 ; p_1 \to p_2 \dashrightarrow p_2$

Beweis. $\quad p_1 \vdash \quad 1)\ p_1$

$p_1 \to p_2 \vdash 1 \to \ 2)\ p_2$

(L_6) $\quad p_1 \to p_2 ;\ p_2 \to p_3 \dashrightarrow p_1 \to p_3$

Beweis. $p_1 \to p_2 \vdash p_1 \to \ 1)\ p_2$

$p_2 \to p_3 \vdash 1 \to \ 2)\ p_3$

(L_7) $\quad p_1 , p_2 \to p ;\ p \to p_3 \dashrightarrow p_1 , p_2 \to p_3$

Beweis. $p_1 , p_2 \to p \vdash p_1 , p_2 \to \ 1)\ p$

$p \to p_3 \vdash \ 1 \ \to \ 2)\ p_3$

(L_8) $\quad p_0 \to p_1 ;\ p_0 \to p_2 ;\ p_1 , p_2 \to p \dashrightarrow p_0 \to p$

Beweis. $\quad p_0 \to p_1 \vdash p_0 \ \to \ 1)\ p_1$

$p_0 \to p_2 \vdash p_0 \ \to \ 2)\ p_2$

$p_1 , p_2 \to p \ \vdash 1, 2 \to \ 3)\ p$

$^{(x)}p$ bezeichne im folgenden stets eine Aussage, in der x nicht frei vorkommt.

(L_9) $\quad {}^{(x)}p_0 \underset{x}{\to} p(x) ;\ p\underset{x}{(x)} \to p \dashrightarrow {}^{(x)}p_0 \to p$

Beweis. $\quad {}^{(x)}p_0 \underset{x}{\to} p(x) \vdash {}^{(x)}p_0 \to \ 1)\ p(x)$

$p\underset{x}{(x)} \to \ p \ \vdash \ \underset{x}{1} \ \to \ 2)\ p$

Käme x in p_0 frei vor, so würde die erste Zeile bei Spezialisierung von x nur dann richtig bleiben, wenn auch p_0 spezialisiert würde.

(L_{10}) $\quad p_1 ;\ p_1 , p_2 \to p \dashrightarrow p_2 \to p$

Beweis. $\quad p_1 \vdash \quad 1)\ p_1$

$p_1 , p_2 \to p \ \vdash 1, p_2 \to \ 2)\ p$

(L_{11}) $\quad p_2 ;\ p_1 , p_2 \to p \dashrightarrow p_1 \to p$

Beweis. Entsprechend zu (L_{10}).

(L_{12}) $\quad p_1 , p_2 \to p \dashrightarrow p_1 \dashrightarrow p_2 \to p$

Beweis. $\quad p_1 , p_2 \to p \vdash \quad 1)\ p_1 , p_2 \to p$

$L_{10} \vdash p_1 , \ 1 \to \ 2) \quad p_2 \to p$

(L_{13}) $\quad p_1 \dashrightarrow p_2 \to p \dashrightarrow p_1 , p_2 \to p$

Beweis. $\quad p_1 \dashrightarrow p_2 \to p \vdash p_1 \ \to \ 1)\ p_2 \to p$

$L_5 \vdash p_2 , \ 1 \to \ 2)\ p$

Gilt $p_1 \to p_2$ und $p_2 \to p_1$, dann schreiben wir zur Abkürzung $p_1 \leftrightarrow p$ ($\leftrightarrow$ gelesen: äquivalent).

L_{12}, L_{13} ergeben also $p_1, p_2 \to p \leftarrow\!\!\cdot\!\!\to p_1 \dot{\to} p_2 \to p$.

Allgemein gilt $p_1, \ldots, p_m,\ p'_1, \ldots, p'_n \to p \leftarrow\!\!\cdot\!\!\to p_1, \ldots, p_m \dot{\to} p'_1, \ldots, p'_n \to p$. Dieser Satz gestattet jede Aussage äquivalent umzuformen in eine solche, in der jeder Bestandteil $p_1, \ldots, p_n \to p$ in p kein $\to$ mehr enthält. (Wird ein Bestandteil äquivalent ersetzt, dann entsteht eine äquivalente Aussage). Die Herstellung dieser „primendigen Normalform" ermöglicht eine leichtere Entscheidung über die konsequenzlogische Beweisbarkeit. Um z. B.

(L_{14}) $\quad p_1 \to p'_1;\ p_2 \to p'_2 \dot{\to} p_3 \to p'_3 \ddot{\to} p,\ p_1 \to p'_1;\ p,\ p_2 \to p'_2 \dot{\to} p,\ p_3 \to p'_3$

zu beweisen, genügt es,

$$p''_1,\ p''_2 \to p''_3 \ddot{\to} p \to p''_1;\ p \to p''_2 \dot{\to} p \to p''_3$$

zu beweisen (und dann p''_i durch $p_i \to p'_i, \ldots$ zu ersetzen). Diese letztere Aussage ergibt sich aus (L_9).

(L_{15}) $\quad p_1 \underset{x}{\to} p'_1 \dot{\to} p_2 \to p'_2 \ddot{\to}^{(x)} p,\ p_1 \underset{x}{\to} p'_1 \dot{\to}^{(x)} p,\ p_2 \to p'_2.$

Hier geht der Beweis über

$$p''_1 \underset{x}{\to} p''_2 \ddot{\to}^{(x)} p \underset{x}{\to} p''_1 \dot{\to}^{(x)} p \to p''_2,$$

was aus (L_9) folgt.

Wegen der Einschränkung unseres Aussagenbegriffes $(P_1 - P_4)$ kann die hier benutzte Äquivalenz von $p_1, p_2 \underset{x}{\to} p$ mit $p_1 \dot{\to}_x p_2 \to p$ jedoch noch nicht angeschrieben werden. Es ist zwar

$$p_1,\ p_2 \underset{x}{\to} p \ddot{\to} p_1 \dot{\to} p_2 \to p$$

eine Aussage, sogar ein Satz, ebenso

$$p_1 \dot{\to}_x p_2 \to p \ddot{\to} p_1,\ p_2 \to p$$

aber

$$p_1,\ p_2 \underset{x}{\to} p \dot{\to} p_1 \dot{\to}_x p_2 \to p$$

ist noch nicht einmal eine Aussage (wegen des x unter der Konklusion).

§ 3.

Definitionen und Inversionen. Das Induktionsprinzip.

1. Wir haben in § 2 dargestellt, wie die Konsequenzen aus gegebenen Regeln abgeleitet werden. Nehmen wir z. B. die Definition der Gleichheitssätze

(D_1) $\quad | = |$

(D_2) $\quad z_1 = z_2 \to z_1 | = z_2 |.$

Durch Angabe von D_1, D_2 allein ist die Gleichheit noch nicht bestimmt, es fehlt noch die Angabe, daß zur Konstruktion von Gleichheiten auch

nur diese Regeln benutzt werden sollen. Weil zur Ableitung von Aussagen $z_1| = z_2|$ nur (D_2) zur Verfügung steht, muß in jeder Ableitung von $z_1| = z_2|$ die letzte Zeile lauten:

$$D_2 \vdash z_1 = z_2 \to n)\ z_2| = z_2|.$$

$z_1 = z_2$ muß also schon vorher abgeleitet sein. Hiernach ist die Benutzung einer Regel $z_1| = z_2| \to p$ eliminierbar bei Benutzung der Regel $z_1 = z_2 \to p$.

Man hat dazu die Zeile

$$z_1| = z_2| \to p \vdash z_1| = z_2| \to n)\ p$$

zu ersetzen durch die Zeile

$$z_1 = z_2 \to p \vdash z_1 = z_2 \to n)\ p.$$

Die Ableitung bleibt bei dieser Ersetzung richtig, weil die Prämisse $z_1 = z_2$ ebenso wie $z_1| = z_2|$ schon zu den abgeleiteten gehört.

Weil also zur Ableitung von Aussagen $z_1| = z_2|$ nur

$$(D_2) \qquad z_1 = z_2 \to z_1| = z_2|$$

vorhanden ist, gilt die Metaregel

$$(*D_2) \qquad z_1 = z_2 \to p \dashrightarrow z_1| = z_2| \to p$$

Aus $(*D_2)$ gewinnt man selbstverständlich sofort

$$z_1| = z_2| \to z_1 = z_2$$

und hieraus läßt sich auch wieder $(*D_2)$ ableiten. $(*D_2)$ ist also äquivalent mit der Umkehrung von (D_2). Wir nennen deswegen $(*D_2)$ die *Inversion* von (D_2).

Zur Ableitung von Aussagen $z_1 + z_2| = z_3$ steht nur (D_7) zur Verfügung. Die Inversion lautet

$$(*D_7) \qquad z_1 + z_2 = z,\ z| = z_3 \to_z p \dashrightarrow z_1 + z_2| = z_3 \to p.$$

Die Begründung ist analog wie oben. Die Regel $z_1 + z_2| = z_3 \to p$ kann eliminiert werden durch $z_1 + z_2 = z,\ z| = z_3 \to_z p$, da in einer Ableitung, in der $z_1 + z_2| = z_3$ abgeleitet vorkommt, auch Prämissen $z_1 + z_2 = z$ und $z| = z_3$ schon vorher abgeleitet vorkommen müssen. Da auf der linken Seite von (D_7) mehrere Prämissen auftreten, außerdem eine Variable z, die rechts nicht vorkommt, darf hier nicht mehr einfach der Pfeil umgekehrt werden.

Ein arithmetisches Beispiel, in dem mehrere Regeln zur Ableitung einer Aussage zur Verfügung stehen, ergibt sich etwa bei Definition der Ordnungsrelation $<$:

$$z_1| = z_2 \to z_1 < z_2$$

$$z_1 < z,\ z| = z_2 \to z_1 < z_2.$$

Ist eine Aussage $z_1 < z_2$ abgeleitet, so ist in der letzten Zeile eine dieser beiden Regeln benutzt. Eine Regel $z_1 < z_2 \to {}^{(z)}p$ kann also jedenfalls eliminiert werden mit Hilfe der Regeln

$$z_1 = z_2 \to {}^{(z)}p$$

$$z_1 < z, z| = z_2 \to {}^{(z)}p$$

so daß als Inversion die Metaregel

$$z_1| = z_2 \to {}^{(z)}p;\ z_1 < z, z| = z_2 \to {}^{(z)}p \dashrightarrow z_1 < z_2 \to {}^{(z)}p$$

entsteht.

Eine allgemeine Formulierung eines „Inversionsprinzips" wäre etwa: Ist ein System von Regeln vorgegeben, so daß zur Ableitung einer Aussage p_0 n u r die Regeln

$$p_1 \to p_0;\ \cdots;\ p_n \to p_0$$

(evtl. mit gebundenen Variablen) benutzt werden können, so gilt für jede Aussage p, in der gewisse Variable nicht frei vorkommen, die Metaregel

$$p_1 \to p;\ \cdots;\ p_n \to p \dashrightarrow p_0 \to p.$$

Hier entsteht die linke Seite durch Ersetzung von p_0 durch p aus dem obigen System der Regeln.

Zur Begründung dieses Prinzips sei noch hinzugefügt, daß es sich aus der Forderung ergibt, die an die Ableitungen von p_0 gestellt wird — nämlich: N u r die Regeln $p_i \to p_0$ $(i = 1, \ldots, n)$ zu benutzen, so daß jede Ableitung von p_0 eine der Prämissen p_i als schon abgeleitet enthalten muß. Statt mit $p_0 \to p$ von p_0 auf p zu schließen, kann also auch von p_i aus mit der Regel $p_i \to p$ — ohne den Umweg über p_0 — auf p geschlossen werden. Da in der Ableitung von p_0 jede der Prämissen p_i benutzt sein kann, müssen alle Regeln $p_i \to p$ zur Verfügung stehen, um $p_0 \to p$ aus jeder Ableitung zu eliminieren.

2. Entgegen einer häufig vertretenen Meinung ist das Induktionsprinzip durchaus nichts spezifisch arithmetisches. Nehmen wir trotzdem die arithmetische Induktion voraus. Der Variabilitätsbereich der Variablen $z, z_1, \ldots$ ist definiert als die Menge der Zeichen, die nach

(Z_1) $\quad |$

(Z_2) $\quad z \to z|$

konstruierbar sind. Eine Ableitung einer Zahl z_0 ist also nichts anderes als eine Reihe

$$\begin{array}{l} | \\ || \\ \vdots \\ z_0 \end{array}$$

wobei jede Zeile aus der vorhergehenden nach (Z_1), (Z_2) hervorgeht (normalerweise wird — außer im ersten Schritt — nur (Z_2) verwendet werden).

Um eine Aussage $p(z_0)$ zu beweisen, kann man daher folgenden Weg versuchen. Man bildet die Reihe

$$\begin{array}{c} p(\mathsf{I}) \\ \vdots \\ p(z_0) \end{array}$$

die aus der Ableitung von z_0 dadurch entsteht, daß jede Zahl z durch die Aussage $p(z)$ ersetzt wird. Da die Ableitung von z_0 nach den Regeln (Z_1), (Z_2) erfolgt, ist die neue Reihe eine Ableitung von $p(z_0)$ nach den Regeln

(1) $$p(\mathsf{I})$$

(2) $$p(z) \to p(z\,\mathsf{I}).$$

$p(z_0)$ ist also ableitbar, wenn die Regeln (1), (2) eliminierbar sind, d. h. es gilt die Metaregel

(I_1) $$p(\mathsf{I});\ p(z) \underset{z}{\to} p(z\,\mathsf{I}) \to p(z).$$

Aus dieser Begründung der arithmetischen Induktion ist ersichtlich, daß hier von den Zahlen nur benutzt wird, daß sie durch ein System von Konstruktionsregeln definiert werden. Führen wir Variable $x, x_1, \ldots$ ein für die Zeichen, die nach einem gegebenen System von Regeln konstruierbar sind, (also z. B. für die Sätze eines Kalküls), so gilt eine entsprechende Induktionsregel. Werden etwa Variable t, t_1, t_2 für „Terme" definiert durch

(T_1) $$z$$

(T_2) $$t_1,\ t_2 \to (t_1 + t_2)$$

(T_3) $$t_1,\ t_2 \to (t_1 \cdot t_2),$$

so sind hiernach, z. B. $((\mathsf{I} + \mathsf{III}) \cdot \mathsf{II})$, $((\mathsf{II} + (\mathsf{II} \cdot \mathsf{I})) + \mathsf{III})$ Terme. Um eine Aussage $p(t)$ abzuleiten, genügen die Regeln

(1) $$p(z)$$

(2) $$p(t_1),\ p(t_2) \to p(t_1 + t_2)$$

(3) $$p(t_1),\ p(t_2) \to p(t_1 \cdot t_2)$$

Die Induktionsregel lautet hier also

(I_2) $$p(z);\ p(t_1),\ p(t_2) \underset{t_1, t_2}{\to} p(t_1 + t_2);\ p(t_1),\ p(t_2) \underset{t_1, t_2}{\to} p(t_1 \cdot t_2) \to p(t).$$

Eine Induktionsregel läßt sich auch ohne Einführung eigener Variablen aufstellen.

Für den Kalkül zur Konstruktion von Gleichheiten $z_1 = z_2$, der definiert ist durch

(D_1) $\quad \mathfrak{l} = \mathfrak{l}$

(D_2) $\quad z_1 = z_2 \to z_1\mathfrak{l} = z_2\mathfrak{l}$

wird aus einer Ableitung

$$\begin{array}{c} \mathfrak{l} = \mathfrak{l} \\ \vdots \\ z_1 = z_2 \end{array}$$

nach (D_1), (D_2) für jede Aussage $p(z_1, z_2)$ die folgende Reihe

$$\begin{array}{c} \mathfrak{l} = \mathfrak{l} \to p(\mathfrak{l}, \mathfrak{l}) \\ \vdots \\ z_1 = z_2 \to p(z_1, z_2) \end{array}$$

eine Ableitung nach den Regeln

(1) $\quad p(\mathfrak{l}, \mathfrak{l})$

(2) $\quad z_1 = z_2 \to p(z_1, z_2) \dashrightarrow z_1\mathfrak{l} = z_2\mathfrak{l} \to p(z_1\mathfrak{l}, z_2\mathfrak{l})$.

Also gilt die Regel

(I_3) $\quad p(\mathfrak{l}, \mathfrak{l});\ z_1 = z_2 \to p(z_1, z_2) \underset{z_1, z_2}{\dashrightarrow} z_1\mathfrak{l} = z_2\mathfrak{l} \to p(z_1\mathfrak{l}, z_2\mathfrak{l}) \dashrightarrow z_1 = z_2 \to p(z_1, z_2)$.

Bei Einführung von Variablen $s, s_1, s_2, \ldots$ für die Sätze $\mathfrak{l} = \mathfrak{l}, \mathfrak{ll} = \mathfrak{ll}, \mathfrak{lll} = \mathfrak{lll}, \ldots$ ließe sich die Induktion folgendermaßen schreiben — wenn $z_1\mathfrak{l} = z_2\mathfrak{l}$ mit s' bezeichnet wird, falls $z_1 = z_2$ mit s bezeichnet ist —

$$p(\mathfrak{l} = \mathfrak{l});\ p(s) \underset{s}{\dashrightarrow} p(s') \dashrightarrow p(s).$$

Mit Hilfe solcher Induktionen werden wir später z. B. beweisen, daß $\mathfrak{l} \neq \mathfrak{l}$ unableitbar ist, d. h. ungleich allen Sätzen (s. § 7).

§ 4.
Positive Logik.

1. Nennen wir zwei Zeichen äquivalent (im weiteren Sinne), wenn jede Aussage bei Ersetzung dieser Zeichen durcheinander in eine äquivalente (im engeren Sinne) Aussage übergeht, so sind z. B. „p_1, p_2" und „p_2, p_1" äquivalent. Um solche Äquivalenzen in einem bequemen Kalkül erfassen zu können, erweitern wir den bisherigen Aussagenbegriff, so daß die Asymmetrie von $\to$ bezügl. der Aussagen $(p_1, p_2 \to p$ ist eine Aussage, $p \to p_1, p_2$ aber nicht; $p(x) \to p$ ist eine Aussage, aber nicht $p \to p(x)$) aufgehoben wird.

Um ganz sicher zu gehen, daß wir bei dieser Erweiterung nichts Beweispflichtiges „hineinmogeln", führen wir außer dem Komma und der Bindung der Variablen ein neues Verknüpfungszeichen $\frown$ (Konjunktion) und einen neuen Operator $\underset{x}{\wedge}$ (Generalisator) ein. Diese sind umgangsprachlich nicht deutlich zu unterscheiden; wir lesen

$$p_1, p_2 : p_1 \text{ mit } p_2$$
$$p_1 \frown p_2 : p_1 \text{ und } p_2$$
$$\underset{x}{p}(x) : p(x) \text{ für jedes } x$$
$$\underset{x}{\wedge} p(x) : \text{für alle } x\, p(x).$$

Diese Übersetzung ist jedoch unwesentlich. Wir definieren den Gebrauch der neuen Zeichen folgendermaßen:

(P_5) Sind p_1, p_2 Aussagen erster Stufe, dann auch $p_1 \frown p_2$.

(P_6) Ist $p(x)$ Aussage erster Stufe, dann auch $\underset{x}{\wedge} p(x)$.

Als Regeln zur Ableitung von Aussagen erster Stufe mit den neuen Zeichen setzen wir fest:

(D_1^0) $$p_1, p_2 \to p_1 \frown p_2$$

(D_2^0) $$\underset{x}{p}(x) \to \underset{x}{\wedge} p(x).$$

Setzen wir ferner fest, daß für Aussagen mit $\frown$ bzw. $\underset{x}{\wedge}$ nur D_1^0 bzw. D_2^0 benutzt werden sollen, so erhalten wir als Inversionen

$(*D_1^0)$ $$p_1, p_2 \to p \mathrel{\dot{\to}} p_1 \frown p_2 \to p$$

$(*D_2^0)$ $$\underset{x}{p}(x) \to p \mathrel{\dot{\to}} \underset{x}{\wedge} p(x) \to p.$$

Die üblichen „Axiome" über Konjunktion und Generalisation ergeben sich jetzt als Konsequenzen.

(S_1^0) $$p_1 \frown p_2 \to p_1$$

Beweis.

$$*D_1^0 \vdash 1)\ p_1, p_2 \to p_1 \mathrel{\dot{\to}} p_1 \frown p_2 \to p_1$$
$$L_5 \vdash L_2, 1 \to S_1^0$$

(S_2^0) $$p_1 \frown p_2 \to p_2$$

Beweis. Entsprechend zu S_1^0.

(S_3^0) $$p_0 \to p_1;\ p_0 \to p_2 \mathrel{\dot{\to}} p_0 \to p_1 \frown p_2$$

Beweis.

$$L_8 \vdash 1)\ p_0 \to p_1;\ p_0 \to p_2;\ p_1, p_2 \to p_1 \frown p_2 \mathrel{\dot{\to}} p_0 \to p_1 \frown p_2$$
$$L_{11} \vdash D_1^0, 1 \to S_3^0$$

(S_4^0) $\quad \bigwedge_x p(x) \to p(x)$

Beweis.

$$*D_2' \vdash 1)\ \underset{x}{p}(x) \to p(x) \dashrightarrow \bigwedge_x p(x) \to p(x)$$

$$L_5 \vdash L_4,\ 1 \to S_4^0.$$

${}^{(x)}p_0$ sei wieder eine Aussage, in der x nicht frei vorkommt.

(S_5^0) $\quad {}^{(x)}p_0 \underset{x}{\to} p(x) \dashrightarrow {}^{(x)}p_0 \to \bigwedge_x p(x)$

Beweis.

$$L_9 \vdash 1)\ {}^{(x)}p_0 \underset{x}{\to} p(x);\ \underset{x}{p}(x) \to \bigwedge_x p(x) \dashrightarrow {}^{(x)}p_0 \to \bigwedge_x p(x)$$

$$L_{11} \vdash D_2^0,\ 1 \to S_5^0.$$

Anschließend läßt sich leicht zeigen, daß aus jedem konsequenzlogischen Satz bei Ersetzung von „," durch „$\frown$" und von $\underset{x_1,\ldots,x_n}{p}(x_1, \ldots, x_n)$ durch $\bigwedge_{z_1} \cdots \bigwedge_{z_2} p(x_1, \ldots, x_n)$ wieder ein Satz entsteht. Es genügt dazu die Äquivalenz von $p_1, \ldots, p_n \to p$ mit $p_1 \frown \cdots \frown p_n \to p$ und von $\underset{x_1,\ldots,x_n}{p}(x_1, \ldots, x_n) \to p$ mit $\bigwedge_{x_1} \cdots \bigwedge_{x_n} p(x_1, \ldots, x_n) \to p$.

Auf die Durchführung im Einzelnen brauchen wir jedoch nicht einzugehen.

2. Nachdem Konjunktion und Generalisation für Aussagen 1. Stufe eingeführt sind, können diese Operationen selbstverständlich auch für die Aussagen höherer Stufe eingeführt werden. Wir erhalten dann eine Menge von Aussagen, die folgendermaßen zu definieren ist:

(P_1') die Primaussagen und -prädikate

(P_2') mit p_1, p_2 auch $p_1 \to p_2$

(P_3') mit p_1, p_2 auch $p_1 \frown p_2$

(P_4') mit $p(x)$ auch $\bigwedge_x p(x)$.

Da jeder Satz von der Form $p_1 \to p_2$ ist, läßt sich eine zweistellige Relation $<$ in der Menge der Aussagen definieren, indem $p_1 < p_2$ gesetzt wird, wenn $p_1 \to p_2$ ein Satz ist. Die Aussagen bilden bezgl. dieser Relation einen „relativorthokomplementären Halbverband" [2]). Bezeichnen wir mit $\vee$ irgendeinen Satz, so ist $\vee$ „obere Spitze". Es gelten nämlich

---

[2]) Diese werden durch die „Axiome" (2)—(8) definiert. Hierbei gehört $\dashrightarrow$ zur Metasprache, während $\to$ in (4), (5) die Operation der relativen Komplementbildung bezeichnet. Ist p_2 eine untere Spitze, so heißt $p_1 \to p_2$ das Orthokomplement von p_1. Vgl. Hermes-Köthe, Enzykl. d. math. Wiss. I 1, 13, 1939.

(1) $p < \vee$ (2) $p < p$ (3) $p_1 < p_2;\ p_2 < p_3 \dot{\to} p_1 < p_3$

(4) $p \frown p_1 < p_2 \dot{\to} p < p_1 \to p_2$ (5) $p < p_1 \to p_2 \dot{\to} p \frown p_1 < p_2$

(6) $p_1 \frown p_2 < p_1$ (7) $p_1 \frown p_2 < p_2$

(8) $p_0 < p_1;\ p_0 < p_2 \dot{\to} p_0 < p_1 \frown p_2$

(9) $\underset{x}{\wedge}\, p(x) < p(x)$

(10) ${}^{(x)}p_0 < p(x) \dot{\to} {}^{(x)}p_0 < \underset{x}{\wedge}\, p(x)$.

(Es handelt sich nur um eine neue Schreibweise für schon bewiesene Sätze).

(1) bis (10) können als „Axiomatisierung" unserer bisherigen Logik gewählt werden. Daß alle konsequenzlogischen Sätze aus diesen Axiomen ableitbar sind, wäre dadurch zu beweisen, daß das Verfahren zum Beweis dieser Sätze in Ableitungen nach (1) bis (10) umgeformt wird. Für die Begründung der Konsequenzlogik ist jedoch durch eine solche Axiomatisierung nichts gewonnen; die Aufgabe einer Begründung besteht ja vielmehr gerade darin, die Herstellung von Aussagen, Sätzen und „Axiomen" als zweckmäßige Handlungen einzuführen.

3. Die Axiomatisierung der bisherigen Logik als Halbverband zeigt, daß immer noch eine unbequeme Unsymmetrie vorhanden ist. Es entsteht die Frage, ob eine Erweiterung des Halbverbandes zu einem Verband möglich ist. Dies wird erreicht durch Einführung zweier weiterer Operationen, der Disjunktion $\smile$ (oder) und der Partikularisation $\underset{x}{\vee}$ (für ein x).

Zunächst wird der Aussagebegriff erweitert durch

(P_5') mit p_1, p_2 auch $p_1 \smile p_2$

(P_6') mit $p(x)$ auch $\underset{x}{\vee}\, p(x)$

und dann der Satzbegriff durch

$(D^0_{3,1})$ $p_1 \to p_1 \smile p_2$

$(D^0_{3,2})$ $p_2 \to p_1 \smile p_2$

$(\dot{D}^0_4)$ $p(x) \to \underset{x}{\vee}\, p(x)$.

Die Einführung dieser Operationen bedarf zu ihrer Begründung nur des Nachweises der Zweckmäßigkeit. Da in den Konsequenzen der Definitionen stets Aussagen mit den „neuen" Zeichen auftreten, wird an der Ableitbarkeit der bisherigen Aussagen nichts geändert. Die Disjunktionen werden sich vor allem dazu als brauchbar erweisen, die Inversion eines beliebigen Systems von Definitionen ohne die Aussagenvariable p zu schreiben: Es wird sich nämlich $p_1 \to p;\ p_2 \to p \underset{p}{\dot{\to}} p_0 \to p$ als äquivalent mit $p_1 \smile p_2 \to p_0$ herausstellen. Die Zweckmäßigkeit dieser

Operationen wird auch unabhängig hiervon schwer zu leugnen sein, man könnte z. B. auf das Auftreten in der Umgangssprache verweisen. Inperhalb der Arithmetik wird $z_1 = z_2 \smile z_1 \neq z_2$ beweisbar. Ohne $\smile$ könnte dieser Satz gar nicht formuliert werden.

Für die neuen Operationen erhalten wir zunächst durch Inversion der Definitionen:

$(*D_3^0)$ $\qquad p_1 \to p_0 ; p_2 \to p_0 \dashv\vdash p_1 \smile p_2 \to p_0$

$(*D_4^0)$ $\qquad p(x) \underset{x}{\to} {}^{(x)}p_0 \dashv\vdash \bigvee_x p(x) \to {}^{(x)}p_0.$

Hiernach ist die Menge der Aussagen jetzt ein Verband. Wir leiten noch einige Sätze über die Distributivität dieses Verbandes ab:

(S_6^0) $\qquad p_0 \frown p_1 \to p ; p_0 \frown p_2 \to p \dashv\vdash p_0 \frown (p_1 \smile p_2) \to p$

Beweis.

$$L_{14} \vdash *D_3^0 \to S_6^0$$

(S_7^0) $\qquad p_0 \frown (p_1 \smile p_2) \to (p_0 \frown p_1) \smile (p_0 \frown p_2)$

Beweis.

$$D_{3,1}^0 \vdash 1)\ p_0 \frown p_1 \to (p_0 \frown p_1) \smile (p_0 \frown p_2)$$
$$D_{3,2}^0 \vdash 2)\ p_0 \frown p_2 \to (p_0 \frown p_1) \smile (p_0 \frown p_2)$$
$$S_6^0 \vdash 1, 2 \to S_7^0.$$

Bekanntlich gilt dann auch die Dualisierung

(S_8^0) $\qquad (p_0 \smile p_1) \frown (p_0 \smile p_2) \to p_0 \smile (p_1 \frown p_2)$

Beweis (abgekürzt).

$$D_{3,2}^0 \vdash \qquad 1)\ p_1 \frown p_2 \to p_0 \smile (p_1 \frown p_2)$$
$$\vdash D_{3,1}^0, S_1^0 \to 2)\ p_0 \frown p_2 \to p_0 \smile (p_1 \frown p_2)$$
$$S_6^0 \vdash 1, 2 \to 3)\ (p_0 \smile p_1) \frown p_2 \to p_0 \smile (p_1 \frown p_2)$$
$$\vdash D_{3,1}^0, S_2^0 \to 4)\ (p_0 \smile p_1) \frown p_0 \to p_0 \smile (p_1 \frown p_2)$$
$$S_6^0 \vdash 3, 4 \to S_8^0.$$

Die „Umkehrungen"

$$(p_0 \frown p_1) \smile (p_0 \frown p_2) \to p_0 \frown (p_1 \smile p_2)$$
$$p_0 \smile (p_1 \frown p_2) \to (p_0 \smile p_1) \frown (p_0 \smile p_2)$$

gelten dagegen in jedem Verband.

Auch für die Partikularisation gilt eine Distributivität.

(S_9^0) $\qquad {}^{(x)}p_0 \frown p(x) \underset{x}{\to} {}^{(x)}p \dashv\vdash {}^{(x)}p_0 \frown \bigvee_x p(x) \to {}^{(x)}p$

Beweis.

$$L_{15} \vdash *D_4^0 \to S_9^0$$

(S^0_{10}) $\qquad {}^{(x)}p_0 \frown \bigvee_x p(x) \to \bigvee_x ({}^{(x)}p_0 \frown p(x))$

Beweis.

$$D^0_4 \vdash 1)\ {}^{(x)}p_0 \frown p(x) \to \bigvee_x ({}^{(x)}p_0 \frown p(x))$$
$$S^0_9 \vdash \tfrac{1}{x} \to S^0_{10}.$$

Die Umkehrung von S^0_{10} ist auch hier noch trivial, dagegen ist die Dualisierung

$$\bigwedge_x ({}^{(x)}p_0 \smile p(x)) \to {}^{(x)}p_0 \smile \bigwedge_x p(x)$$

unableitbar[3]). Die völlige Dualität wird erst später durch die Fiktion des „tertium non datur" hergestellt werden — innerhalb der positiven Logik ist kein Grund vorhanden, die Dualität zu erwarten, denn die Einführungen von Generalisation und Partikularisation sind nicht dual zueinander.

§ 5.

Arithmetik.

Zur Einführung der konsequenzlogischen Symbolik hatten wir in § 1 als Beispiele von Konstruktionsregeln einige arithmetische Definitionen (D_1)—(D_9) eingeführt.

Das Inversionsprinzip liefert sofort darüber hinaus

$(*D_2)$ $\qquad z_1 \mid = z_2 \mid \to z_1 = z_2$

$(*D_3)$ $\qquad z_1 \mid \neq z_2 \mid \to z_1 \neq z_2$

$(*D_6)$ $\qquad z_1 + 1 \mid = z_3 \to z_1 \mid = z_3$

$(*D_7)$ $\qquad z_1 + z_2 \mid = z_3 \to \bigvee_z (z_1 + z_2 = z \frown z \mid = z_3)$

$(*D_8)$ $\qquad z_1 \cdot 1 \mid = z_3 \to z_1 = z_3$

$(*D_9)$ $\qquad z_1 \cdot z_2 \mid = z_3 \to \bigvee_z (z_1 \cdot z_2 = z \frown z + z_1 = z_3).$

Zur Ableitung arithmetischer Sätze haben wir ferner Induktionsregeln für Induktionen über alle Zahlen, alle Terme, alle Sätze usw.

Wir gebrauchen außerdem im folgenden eine Induktion über alle Zahlpaare (z_1, z_2).

(S_1) $\qquad p(1, 1);\ \underset{z}{p}(z \mid, 1);\ \underset{z}{p}(1, z \mid);\ \underset{z_1, z_2}{p(z_1, z_2) \to p(z_1 \mid, z_2 \mid)} \to p(z_1, z_2).$

Der Beweis braucht nur skizziert zu werden. Er läuft über folgende Schritte:

$\vdash$ $\qquad 1)\ p(1, 1) \to p(1, 1)$

$\vdash$ $\qquad 2)\ \underset{z}{p}(z \mid, 1) \to p(z_1 \mid, 1)$

[3]) Zum Beweis der Unableitbarkeit, von der wir keinen weiteren Gebrauch machen werden, vgl. S. C. Kleene, Journal Symb. Logic 10, 1945.

$I_1 \vdash 1, \underset{z}{2} \to 3)\ p(\mathfrak{l}, \mathfrak{l});\ \underset{z}{p}(z\mathfrak{l}, \mathfrak{l}) \to \underset{z_1}{\wedge}\ p(z_1, \mathfrak{l})$

$\vdash \quad 4)\ \underset{z}{p}(\mathfrak{l}, z\mathfrak{l}) \dashrightarrow \underset{z_1}{\wedge}\ p(z_1, z_2) \to p(\mathfrak{l}, z_2\mathfrak{l})$

$\vdash \quad 5)\ p(z_1, z_2) \underset{z_1, z_2}{\to} p(z_1\mathfrak{l}, z_2\mathfrak{l}) \dashrightarrow \underset{z_1}{\wedge}\ p(z_1, z_2) \to p(z_1\mathfrak{l}, z_2\mathfrak{l})$

$I_1 \vdash 4, \underset{z_1}{5} \to 6)\ \underset{z}{p}(\mathfrak{l}, z\mathfrak{l});\ p(z_1, z_2) \underset{z_1, z_2}{\to} p(z_1\mathfrak{l}, z_2\mathfrak{l}) \dashrightarrow \underset{z_1}{\wedge}\ p(z_1, z_2) \to \underset{z_1}{\wedge}\ p(z_1, z_2\mathfrak{l})$

$I_1 \vdash 3, \underset{z_2}{6} \to S_1.$

Mit Hilfe dieser Induktion (S_1) wird der Beweis des wichtigen Satzes

(S_2) $\qquad z_1 = z_2 \smile z_1 \neq z_2$

trivial. Wegen (S_1) genügen die folgenden Zeilen:

$D_1 \dashrightarrow 1)\ \mathfrak{l} = \mathfrak{l} \smile \mathfrak{l} \neq \mathfrak{l}$

$D_3 \to 2)\ z\mathfrak{l} = \mathfrak{l} \smile z\mathfrak{l} \neq \mathfrak{l}$

$D_4 \to 3)\ \mathfrak{l} = z\mathfrak{l} \smile \mathfrak{l} \neq z\mathfrak{l}$

$D_2, D_5 \to 4)\ z_2 = z_2 \smile z_1 \neq z_2 \to z_1\mathfrak{l} = z_2\mathfrak{l} \smile z_1\mathfrak{l} \neq z_2\mathfrak{l}.$

Hiernach gilt für jedes z_1, z_2 mindestens eine der Aussagen $z_1 = z_2$, $z_1 \neq z_2$. Daß auch nur höchstens eine gilt, können wir jetzt noch nicht beweisen; dies wird erst nach Einführung einer Negation möglich.

Für die Gleichheit beweisen wir als nächstes die sog. Identitätsregeln.

(S_3) $\qquad z = z.$

Beweis nach I_1.

$D_1 \to 1)\ \mathfrak{l} = \mathfrak{l}$

$D_2 \to 2)\ z = z \to z\mathfrak{l} = z\mathfrak{l}.$

(S_4) $\qquad z_1 = z_2,\ p(z_1) \to p(z_2).$

Wir beweisen hier nach (I_3) die äquivalente Aussage

$$z_1 = z_2 \to \underset{p}{\wedge}\ (p(z_1) \to p(z_2)).$$

Es genügt hierfür

1) $p(\mathfrak{l}) \to p(\mathfrak{l})$

2) $z_1 = z_2 \to \underset{p}{\wedge}\ (p(z_1) \to p(z_2)) \dashrightarrow z_1 = z_2 \dashrightarrow p(z_1\mathfrak{l}) \to p(z_2\mathfrak{l})$

$^*D_2, \underset{p}{2} \to 3)\ z_1 = z_2 \to \underset{p}{\wedge}\ (p(z_1) \to p(z_2)) \dashrightarrow z_1\mathfrak{l} = z_2\mathfrak{l} \to \underset{p}{\wedge}\ (p(z_1\mathfrak{l}) \to p(z_2\mathfrak{l}))$

$I_3 \vdash 1, \underset{z_1, z_2}{3} \to S_4.$

Die Benutzung von gebundenen Aussagevariablen ist hier genau so unbedenklich wie der Gebrauch von Zahlvariablen, da ja der Variabilitätsbereich konstruktiv definiert ist. Als Konsequenzen von (S_3), (S_4) erhält man die üblichen Gleichheitsaxiome

$$z_1 = z_3,\ z_2 = z_3 \to z_1 = z_2$$

bzw.

$$z_1 = z_2 \to z_2 = z_1$$

$$z_1 = z_2,\ z_2 = z_3 \to z_1 = z_3.$$

Der Ausbau der Arithmetik geschieht jetzt durch Hinzunahme „rekursiver Funktionen". Wir behandeln hier als Beispiel nur Addition und Multiplikation. Diese werden zunächst eingeführt durch (D_8)—(D_9) mit den dazugehörigen Inversionen $(*D_8)$—$(*D_9)$. Dann werden Terme eingeführt gemäß (T_1)—(T_3) in § 3. Unter den Aussagen $t_1 = t_2$, $t_1 \neq t_2$ über Terme werden als Sätze diejenigen definiert, die nach

(D_{10}) $\quad t_1 = z_1,\ t_2 = z_2,\ z_1 + z_2 = z \to t_1 + t_2 = z$

(D_{11}) $\quad t_1 = z_1,\ t_2 = z_2,\ z_1 \cdot z_2 = z \to t_1 \cdot t_2 = z$

(D_{12}) $\quad t_1 = z_1,\ t_2 = z_2, \quad z_1 = z_2 \to t_1 = t_2$

(D_{13}) $\quad t_1 = z_1,\ t_2 = z_2, \quad z_1 \neq z_2 \to t_1 \neq t_2$

ableitbar sind. Da jede Zahl nach (T_1) ein Term ist, sind (D_{10})—(D_{13}) nur insoweit Definitionen, als die auftretenden Terme keine Zahlen sind. Um (D_{10})—(D_{13}) allgemein zulassen zu können, müssen

(S_5) $\quad z_1' = z_1,\ z_2' = z_2,\ z_1 + z_2 = z \to z_1' + z_2' = z$

(S_6) $\quad z_1' = z_1,\ z_2' = z_2,\ z_1 \cdot z_2 = z \to z_1' \cdot z_2' = z$

(S_7) $\quad z_1' = z_1,\ z_2' = z_2, \quad z_1 = z_2 \to z_1' = z_2'$

(S_8) $\quad z_1' = z_1,\ z_2' = z_2, \quad z_1 \neq z_2 \to z_1' \neq z_2'$

bewiesen werden. Diese Sätze ergeben sich aber leicht aus den Identitätsregeln (S_3), (S_4).

Als Inversionen gehören zu (D_{10})—(D_{13})

$(*D_{10})$ $\quad t_1 + t_2 = z \to \bigvee_{z_1}\bigvee_{z_2}(t_1 = z_1 \wedge t_2 = z_2 \wedge z_1 + z_2 = z)$

$(*D_{11})$ $\quad t_1 \cdot t_2 = z \to \bigvee_{z_1}\bigvee_{z_2}(t_1 = z_1 \wedge t_2 = z_2 \wedge z_1 \cdot z_2 = z)$

$(*D_{12})$ $\quad t_1 = t_2 \to \bigvee_{z_1}\bigvee_{z_2}(t_1 = z_1 \wedge t_2 = z_2 \wedge z_1 = z_2)$

$(*D_{13})$ $\quad t_1 \neq t_2 \to \bigvee_{z_1}\bigvee_{z_2}(t_1 = z_1 \wedge t_2 = z_2 \wedge z_1 \neq z_2).$

Bisher ist die Benutzung des Gleichheitszeichens in den Aussagen $z_1 + z_2 = z_3$, $z_1 \cdot z_2 = z_3$ nicht begründet. Sie empfiehlt sich jedoch, da

die folgenden Sätze gelten:

(S_9) $\quad z_1+z_2=z_3',\ z_1+z_3=z_3''\to z_3'=z_3''$

(S_{10}) $\quad z_1\cdot z_2=z_3',\ z_1\cdot z_3=z_3''\to z_3'=z_3''.$

Beweis von (S_9): Induktion über z_2.

$*D_6\to 1)\ z_1+|=z_3'\to z_1|=z_3'$

$*D_6\to 2)\ z_1+|=z_3''\to z_1|=z_3''$

$S_4, 1, 2\to 3)\ z_1+|=z_3',\ z_1+|=z_3''\to z_3'=z_3''$

$*D_7\to 4)\ z_1+z_2|=z_3'\to \underset{z'}{\vee}(z_1+z_2=z'\wedge z'|=z_3')$

$*D_7\to 5)\ z_1+z_2|=z_3''\to \underset{z''}{\vee}(z_1+z_2=z''\wedge z''|=z_3'')$

$S_4, 4, 5\to 6)\ z_1+z_2|=z',\ z_1+z_2=z''\to z'=z''\to z_1+z_2|=z_3',\ z_1+z_2|=z_3''\to z_3'=z_3''$

$I_1\vdash \underset{z_2}{3, 6}\to S_9.$

Der Beweis von (S_{10}) ist jetzt mit Hilfe von (S_9) analog zu führen.

Wegen (S_9), (S_{10}) liefern die Prädikate $z_1+z_2=z_3$, $z_1\cdot z_2=z_3$ „Funktionen", nämlich z_3 als Funktion von (z_1, z_2).

Daß für jedes Paar (z_1, z_2) ein Funktionswert existiert, wird ausgesagt durch

(S_{11}) $\quad \underset{z}{\vee}\ z_1+z_2=z$

(S_{12}) $\quad \underset{z}{\vee}\ z_1\cdot z_2=z.$

Beweis von (S_{11}): Induktion über z_2.

$1)\ z_1|=z_1|\to \underset{z}{\vee}\ z_1|=z$

$S_3, 1\to 2)\ \underset{z}{\vee}\ z_1|=z$

$D_6, 2\to 3)\ \underset{z}{\vee}\ z_1+|=z$

$\underset{z_1}{2}\to 4)\ \underset{z'}{\vee}\ z|=z'$

$D_7, 4\to 5)\ \underset{z}{\vee}\ z_1+z_2=z\to \underset{z'}{\vee}\ z_1+z_2|=z'$

$I_1\vdash \underset{z_2}{3, 5}\to S_{11}.$

Mit (S_{11}) ist der Beweis von (S_{12}) analog zu führen.

Durch Induktion über alle Terme gewinnen wir aus (S_9), (S_{10}) den allgemeineren Satz

(S_{13}) $\quad t=z',\ t=z''\to z'=z''.$

Der Beweis sei nur skizziert. Für Terme, die Zahlen sind, folgt (S_{13}) aus (S_4). Gilt (S_{13}) für t_1, t_2, so folgt (S_{13}) für $t_1 + t_2$, denn aus

$$t_1 + t_2 = z', \quad t_1 + t_2 = z''$$

folgt

$$\bigvee_{z_1'} \bigvee_{z_2'} (t_1 = z_1' \wedge t_2 = z_2' \wedge z_1' + z_2' = z'), \quad \bigvee_{z_1''} \bigvee_{z_2''} (t_1 = z_1'' \wedge t_2 = z_2'' \wedge z_1'' + z_2'' = z'').$$

Hieraus folgt weiter

$$z_1' = z_1'', \quad z_2' = z_2''$$

und dann mit (S_4) und (S_9)

$$z' = z'',$$

analog folgt (S_{13}) für

$$t_1 \cdot t_2.$$

Auch (S_{11}), (S_{12}) sind zu verallgemeinern zu

$$(S_{14}) \qquad \bigvee_z t = z.$$

Beweis ebenfalls durch Induktion über alle Terme!

Nach (S_{13}), (S_{14}) ist jeder Ausdruck, der wie ein Term gebaut ist, aber Zahlvariable enthält, eine Funktion, d. h. überall definiert und eindeutig. Die Definition (D_{12}) der Sätze $t_1 = t_2$ zusammen mit $(*D_{12})$ und den Sätzen (S_{13}), (S_{14}) liefert die Gleichheitsaxiome für Terme — und damit die Begründung des Gebrauches des Gleichheitszeichens.

$$(S_{15}) \qquad t = t.$$

Beweis.

$$S_3, S_{14} \to 1) \bigvee_{z_1} \bigvee_{z_2} (t = z_1 \wedge t = z_2 \wedge z_1 = z_2)$$

$$D_{12}, 1 \to S_{15}.$$

$$(S_{16}) \qquad t_1 = t_2, t_2 = t_3 \to t_1 = t_2.$$

Beweis.

$$*D_{12} \to 1)\ t_1 = t_2 \to \bigvee_{z_1} \bigvee_{z_1'} (t_1 = z_1 \wedge t_2 = z_1' \wedge z_1 = z_1')$$

$$*D_{12} \to 2)\ t = t_2 \to \bigvee_{z_2} \bigvee_{z_2'} (t_2 = z_2 \wedge t_2 = z_2' \wedge z_2 = z_2')$$

$$S_{12} \to 3)\ t_2 = z_1', t = z_2' \to z_1' = z_2'$$

$$S_4 \to 4)\ z_1 = z_1, z_2 = z_2', z_1' = z_2' \to z_1 = z_2$$

$$D_{12} \to 5)\ t_1 = z_1, t_2 = z_2, z_1 = z_2 \to t_1 = t_2$$

$$1, \ldots, 5 \to S_{16}.$$

Auch die allgemeinere Identitätsregel

$$t_1 = t_2, \quad p(t_1) \to p(t_2)$$

ist leicht zu beweisen, so daß wir nicht näher darauf einzugehen brauchen.

Mit Hilfe von (D_{12}), (D_{13}) und (S_{14}) überträgt sich auch noch das tertium non datur (S_9) auf Terme.

(S_{17}) $$t_1 = t_2 \smile t_1 \neq t_2 .$$

Beweis.

$$S_{14} \rightarrow 1) \bigvee_{z_1} \bigvee_{z_2} (t_1 = z_1 \frown t_2 = z_2)$$
$$S_9 \rightarrow 2) z_1 = z_2 \smile z_1 \neq z_2$$
$$D_{12} \rightarrow 3) z_1 = z_2, t_1 = z_1, t_2 = z_2 \rightarrow t_1 = t_2$$
$$D_{13} \rightarrow 4) z_1 \neq z_2, t_1 = z_1, t_2 = z_2 \rightarrow t_1 \neq t_2$$
$$1, \ldots, 4 \rightarrow S_{17}.$$

Nach dem Muster der hier behandelten Addition und Multiplikation können weitere rekursive Funktionen hinzugenommen werden. Ein allgemeines Schema zur Einführung einer Funktion $f(z)$ ist etwa

(1) $$p(z_2) \rightarrow f(|) = z_2$$

(2) $$f(z_1) = z, \quad p(z, z_1, z_2) \rightarrow f(z_1|) = z_2,$$

worin p Prädikate mit

$$p(z_2'), p(z_2'') \rightarrow z_2' = z_2'', \quad p(z, z_1, z_2'), p(z, z_1, z_2'') \rightarrow z_2' = z_2''$$

sind. Die Inversionen lauten:

(*1) $$f(|) = z_2 \rightarrow p(z_2)$$

(*2) $$f(z_1|) = z_2 \rightarrow \bigvee_z (f(z_1) = z \frown p(z, z_1, z_2)).$$

Aus ihnen folgt wie oben

$$f(z_1) = z_2', f(z_1) = z_2'' \rightarrow z_2' = z_2''.$$

Damit auch

$$\bigvee_z f(z_1) = z$$

gilt, müssen die Prädikate p entsprechende Bedingungen erfüllen.

Definiert man für zwei solche Funktionen f_1, f_2 die Aussagen

$$f_1(z_1) = f_2(z_2)$$
$$f_1(z_1) \neq f_2(z_2)$$

entsprechend zu (D_{14}), (D_{15}), so wird ebenfalls das tertium non datur übertragbar.

Alle in der Arithmetik üblicherweise vorkommenden Primprädikate lassen sich mit Hilfe solcher rekursiver Funktionen definieren, z. B. ist die Einführung von $z_1 < z_2$ überflüssig, da die Addition das Prädikat $\bigvee_z z_1 + z = z_2$ liefert.

§ 6.

Semiotik.

Um im nächsten § die „Unableitbarkeit" einer Aussage p dadurch definieren zu können, daß jeder Satz „ungleich" p ist, wollen wir zunächst Gleichheit und Ungleichheit von den Zahlen auf beliebige Zeichen ausdehnen. Die so entstehende Semiotik wird also nichts anderes als eine verallgemeinerte Arithmetik sein.

Wir gehen von endlich vielen Atomzeichen aus, z. B.

$$\mathsf{I},\ =,\ \neq,\ +,\ \cdot,\ (,\ ),\ z,\ \to,\ \frown,\ \smile,\ \underset{z}{\wedge},\ \underset{z}{\vee}.$$

Statt z lassen wir auch endlich viele andere Zeichen, z. B. $z_1, \ldots, z_n$ zu — für jede arithmetische Untersuchung hat man sich also zunächst mit genügend Variablen zu versorgen.

Als Variable für Atomzeichen gebrauchen wir a, $a_1, a_2, \ldots$.

Als Variable für beliebige Zeichen, die aus diesen Atomen zusammengesetzt werden können, benutzen wir $x, x_1, x_2, \ldots$.

Der Variabilitätsbereich der x wird also definiert durch

(X_1) $\qquad a$

(X_2) $\qquad x \to x\,a.$

Hiernach gilt eine Induktionsregel:

(I_1^{00}) $\qquad p(\underset{a}{a});\ p(x) \underset{x,\,a}{\to} p(x\,a) \to p(x).$

Analog zur Arithmetik läßt sich hieraus leicht eine Induktion über Paare gewinnen:

$$(S_1^{00})\ p(\underset{a_1,\,a_2}{a_1, a_2});\ p(\underset{x,\,a_1,\,a_2}{x a_1, a_2});\ p(\underset{x,\,a_1,\,a_2}{a_1, x a_2});\ p(x_1, x_2) \underset{x_1, x_2, a_1, a_2}{\to} p(x_1 a_1, x_2 a_2) \to p(x_1, x_2).$$

Für Atomzeichen definieren wir Gleichheit und Ungleichheit (wir schreiben $\equiv$ und $\not\equiv$ der Übersichtlichkeit wegen) durch Aufzählung in einer Tabelle:

	I	$=$	$\neq$	$\cdots$	$\underset{z}{\vee}$
I	$\mathsf{I} \equiv \mathsf{I}$	$\mathsf{I} \not\equiv =$	$\mathsf{I} \not\equiv \neq$	$\cdots$	$\mathsf{I} \not\equiv \underset{z}{\vee}$
$=$	$= \not\equiv \mathsf{I}$	$= \equiv =$			
$\neq$	$\neq \not\equiv \mathsf{I}$				
$\vdots$	$\vdots$			$\ddots$	
$\underset{z}{\vee}$	$\underset{z}{\vee} \not\equiv \mathsf{I}$				$\underset{z}{\vee} \equiv \underset{z}{\vee}$

Nur in der Hauptdiagonale tritt $\equiv$ auf, sonst stets $\not\equiv$.

Die Regel

$$(S_2^{00}) \qquad a \equiv a$$

ist daher eliminierbar — umgekehrt folgen hieraus alle Sätze der Hauptdiagonale. Für die Ungleichheit ist die Tabelle dagegen nicht zu vermeiden.

Für zusammengesetzte Zeichen definieren wir

$$(D_1^{00}) \qquad x_1 \equiv x_2, \quad a_1 \equiv a_2 \to x_1 a_1 \equiv x_2 a_2$$

$$(D_2^{00}) \qquad x a_1 \not\equiv a_2$$

$$(D_3^{00}) \qquad a_1 \not\equiv x a_2$$

$$(D_{4,1}^{00}) \qquad x_1 \not\equiv x_2 \to x_1 a_1 \not\equiv x_2 a_2$$

$$(D_{4,2}^{00}) \qquad a_1 \not\equiv a_2 \to x_1 a_1 \not\equiv x_2 a_2.$$

Hierzu gehören als Inversionen

$$(*D_1^{00}) \qquad x_1 a_1 \equiv x_2 a_2 \to x_1 \equiv x_2 \frown a_1 \equiv a_2$$

$$(*D_4^{00}) \qquad x_1 a_1 \not\equiv x_2 a_2 \to x_1 \not\equiv x_2 \smile a_1 \not\equiv a_2.$$

Die Identitätsregeln

$$(S_3^{00}) \qquad x \equiv x$$

$$(S_4^{00}) \qquad x_1 \equiv x_2, \quad p(x_1) \to p(x_2)$$

ergeben sich analog zur Arithmetik (es muß nur zunächst

$$a_1 \equiv a_2, \quad p(a_1) \to p(a_2)$$

bewiesen werden. Dies geschieht durch Rückgang auf die Tabelle.) Die hier zu benutzende Regel für Induktionen über alle Gleichheitssätze lautet

$$(I_2^{00}) \qquad p(\underset{a}{a}, a);$$

$$x \equiv x_2 \to p(x\ , x_2)_{x_1, x_2, \overrightarrow{a_1, a_1}} x_1 a_1 \equiv x_2 a_2 \to p(x_1 a_1, x_2 a_2) \dashrightarrow x_1 \equiv x_2 \to p(x_1 x_2).$$

Während in der Arithmetik $|z = z|$ gilt (durch Induktion über z beweisbar), gilt in der Semiotik $ax \equiv xa$ nicht allgemein. Es ist daher nützlich, die einseitige Bevorzugung der Operation xa gegenüber ax aufzuheben.

$$(S_5^{00}) \qquad x_1 \equiv x_2, \quad a_1 \equiv a_2 \to a\ x_1 \equiv a_2 x_2.$$

Beweis trivial mithilfe von (I_2^{00}).

(S_6^{00}) $\qquad a_1 x \neq a_2$

(S_7^{00}) $\qquad a_1 \neq x a_2 .$

Beweise trivial mit I_1^{00} über x.

(S_8^{00}) $\qquad x_1 \neq x_2 \to a_1 x_1 \neq a_2 x_2$

(S_9^{00}) $\qquad a_1 \neq a_2 \to a\, x_1 \neq a_2 x_2 .$

Beweise durch Induktion über $(x\ , x_2)$ mit (S_6^{00}), (S_7^{00}).

Im Unterschied zur Arithmetik ist für die Semiotik ferner die Relation „a kommt in x vor" wichtig. Wir definieren zwei Prädikate $a \iota x$ und $a \bar{\iota} x$ unabhängig voneinander durch

(D_5^{00}) $\qquad a_1 \equiv a_2 \to a_1 \iota a_2$

$(D_{6,1}^{00})$ $\qquad a_1 \iota x \to a_1 \iota x a_2$

$(D_{6,2}^{00})$ $\qquad a_1 \iota a_2 \to a_1 \iota x a_2$

(D_7^{00}) $\qquad a_1 \neq a_2 \to a_1 \bar{\iota}\, a_2 .$

(D_8^{00}) $\qquad a_1 \bar{\iota}\, x, \quad a_1 \bar{\iota}\, a_2 \to a_1 \bar{\iota}\, x a_2 .$

Als Inversionen gehören hierzu

$(*D_5^{00})$ $\qquad a_1 \iota a_2 \to a_1 \equiv a_2$

$(*D_6^{00})$ $\qquad a_1 \iota x a_2 \to a_1 \iota x \vee a_1 \iota a_2$

$(*D_7^{00})$ $\qquad a_1 \bar{\iota}\, a_2 \to a_1 \neq a_2$

$(*D_8^{00})$ $\qquad a_1 \bar{\iota}\, x a_2 \to a_1 \bar{\iota}\, x \wedge a_1 \bar{\iota}\, a_2 .$

Wir brauchen über diese Prädikate vor allem

(S_{10}^{00}) $\qquad a \iota x_1, \quad a \bar{\iota}\, x_2 \to x_1 \neq x_2 .$

Beweis durch Induktion über (x_1, x_2)!

Dieser Satz liefert in vielen Fällen das einfachste Mittel, um $x_1 \neq x_2$ zu beweisen.

Da wir hier die Semiotik nur für den Zweck von Unableitbarkeitsbeweisen gebrauchen, gehen wir auf den weiteren Ausbau (z. B. Einführung von Verkettung und Ersetzung, analog zu Addition und Multiplikation) nicht ein.

§ 7.

Unableitbarkeit.

Wir beschränken uns auf die Unableitbarkeit bezügl. des arithmetischen Kalküls, der durch $(D_1) - (D_{13})$, $(D_1^0) - (D_4^0)$ (einschließlich Definition der Variabilitätsbereiche (Z_1), (Z_2), $(T_1) - (T_3)$) gegeben ist.

Für die Frage, ob eine Aussage p nach diesen Regeln einschließlich der hinzugenommenen Inversionen, Induktionen und der bewiesenen Sätze ableitbar ist, brauchen wir nur die Definitionen selbst zu berücksichtigen. Denn der Beweis der hinzugenommenen Regeln bestand ja gerade immer darin, zu zeigen, daß diese Regeln aus jeder Ableitung eliminierbar seien.

Wir nennen eine Aussage p unableitbar, wenn $p \overset{s}{\nparallel} s$ für jeden arithmetischen Satz s gilt, formalisiert

$$(D_{14}) \qquad p \overset{s}{\nparallel} s \rightarrow \neg p.$$

Hierzu gehört als Inversion

$$(*D_{14}) \qquad \neg p \rightarrow \bigwedge_{s} p \overset{s}{\nparallel} s.$$

Der Variabilitätsbereich von s ist hierbei definiert durch $(D_1) - (D_{13})$, $(D_1^0) - (D_4^0)$. Die Induktionsregel für Induktionen über s — wir nennen sie Satzinduktionen — explizit hinzuschreiben, ist nicht erforderlich, da sie für jeden einzelnen Fall leicht aufzustellen ist. Um die Unableitbarkeit von $| \neq |$ zu zeigen, hat man $| \neq | \overset{s}{\nparallel} s$ durch Satzinduktion zu beweisen. Von den Definitionen $(D_1) - (D_{13})$, $(D_1^0) - (D_4^0)$ brauchen dabei nur $(D_3) - (D_5)$ berücksichtigt zu werden, da die Konklusionen der anderen Definitionen trivialerweise $\overset{s}{\nparallel} | \neq |$ sind: es kommt in ihnen stets ein Atomzeichen a vor mit $a\,\bar{\imath}\,| \neq |$, (s. (S_{10}^{00})). Dies gilt auch für (D_{13}), außer wenn t_1, t_2 Zahlen sind. Für diesen Fall aber haben wir (D_{13}) als eliminierbar bewiesen.

$$(D_3) \qquad z\,| \neq |\,\overset{s}{\nparallel}\,| \neq |.$$

Beweis.

$$D_2^{00} \rightarrow 1)\ z\,|\,\overset{s}{\nparallel}\,|$$
$$D_{4,1}^{00},\ 1 \rightarrow 2)\ z\,| \neq |\,\overset{s}{\nparallel}\,| \neq |.$$

$$(D_4) \qquad | \neq z\,|\,\overset{s}{\nparallel}\,| \neq |.$$

Beweis.

$$D_2^{00} \rightarrow 1)\ z\,|\,\overset{s}{\nparallel}\,|$$
$$S_8^{00},\ 1 \rightarrow 2)\ | \neq z\,|\,\overset{s}{\nparallel}\,| \neq |.$$

$$(D_5) \qquad z_1\,| \neq z_2\,|\,\overset{s}{\nparallel}\,| \neq |.$$

Beweis.

$$D_{4,2}^{00},\ |\,\overset{s}{\nparallel} \neq \rightarrow 1)\ z_1\,| \neq z_2\,\overset{s}{\nparallel}\,| \neq$$
$$D_{4,1}^{00},\ 1 \rightarrow 2)\ z_1\,| \neq z_2\,|\,\overset{s}{\nparallel}\,| \neq |.$$

Damit ist die Unableitbarkeit von $| \neq |$ schon bewiesen. Für die Unableitbarkeit von $z\,| = |$ braucht entsprechend nur (D_1), (D_2) berücksichtigt zu werden, da in den Konklusionen der anderen Definitionen wieder Atomzeichen vorkommen, die nicht in $z\,| = |$ vorkommen.

(D_1) $$\mid = \mid \ddagger z \mid = \mid.$$

Beweis.

$$D_2^{\infty} \rightarrow 1) \mid \ddagger z \mid$$
$$D_{4,1}^{\infty},\ 1 \rightarrow 2) \mid = \mid \ddagger z \mid = \mid.$$

(D_2) $$z_1 \mid = z_2 \mid \ddagger z \mid = \mid.$$

Beweis.

$$D_{4,2}^{\infty\infty},\ \mid \ddagger = \rightarrow 1)\ z_1 \mid = z_2 \ddagger z \mid =$$
$$D_{4,1}^{\infty},\ 1 \rightarrow 2)\ z_1 \mid = z_2 \mid \ddagger z \mid = \mid.$$

Als letztes beweisen wir noch die Unableitbarkeit von $\mid = z \mid$.

(D_1) $$\mid = \mid \ddagger \mid = z \mid.$$

Beweis.

$$D_2^{\infty} \rightarrow 1) \mid \ddagger z \mid$$
$$S_8^{\infty},\ 1 \rightarrow 2) \mid = \mid \rightarrow z \mid = \mid.$$

(D_2) $$z_1 \mid = z_2 \mid \ddagger \mid = z \mid.$$

Beweis.

$$S_9^{\circ},\ \mid \ddagger = \rightarrow 1)\ z_1 = z_2 \mid \ddagger = z \mid$$
$$S_8^{\infty},\ 1 \rightarrow 2) \mid z_1 = z_2 \mid \ddagger \mid = z \mid$$
$$S_4^{\infty} \mid\!\!- \mid z_1 \equiv z_1 \mid,\ 2 \rightarrow 3)\ z_1 \mid = z_2 \mid \ddagger \mid = z. \mid.$$

Für kompliziertere Aussagen (z. B. schon für $z_1 = z_2 \frown z_1 \neq z_2$) wird diese direkte Methode der Unableitbarheitsbeweise unbrauchbar. Wir entwickeln deshalb im nächsten § eine indirekte Methode, um die Unableitbarkeit einer Aussage zu gewinnen.

§ 8.

Negation.

Sind p_1, p_2 ableitbare Formeln, so gilt die Äquivalenz

$$p_1 \leftrightarrow p_2.$$

Diese gestattet es, ein Zeichen $\vee$ (das Wahre) einzuführen, mit der Festsetzung, daß $\vee$ mit jeder ableitbaren Aussage äquivalent ist.

Um ein entsprechendes Zeichen $\wedge$ (das Falsche) mit der Bestimmung, daß, $\wedge$ mit jeder unableitbaren Aussage äquivalent ist, einführen zu können, setzen wir daher zunächst fest, daß unableitbare Aussagen untereinander äquivalent sein sollen, formalisiert:

$$\neg p_1,\ \neg p_2 \rightarrow p_1 \rightarrow p_2.$$

Da sich bekanntlich (nach B. Russell) „Festsetzung" und „Beweis" zueinander verhalten, wie „Diebstahl" und „ehrlicher Erwerb", muß diese Festsetzung begründet werden — also gleichsam als „ehrlicher

Diebstahl" nachgewiesen werden. Es ist zu zeigen, daß die Regel $p_1 \to p_2$ aus jeder Ableitung eliminiert werden kann, falls p_1, p_2 unableitbar sind. Unter dieser Voraussetzung kommt aber $p_1 \to p_2$ in keiner Ableitung vor, denn p_1 kann nicht als schon abgeleitete Prämisse vorkommen. Diese Begründung würde auch für die Metaregel

$$\neg p_1 \dashrightarrow p_1 \to p_2$$

hinreichen, wir werden jedoch mit der oben geforderten Äquivalenz aller unableitbaren Aussagen auskommen.

Zunächst können wir jetzt $\wedge$ (das Falsche) einführen. $\wedge$ kann für jede der Aussagen $| \neq |$, $z | = |$, $| = | z$ stehen:

$$\neg p \dashrightarrow p \leftrightarrow \wedge .$$

Anschließend definieren wir eine Negation, indem wir $\bar{p}$ als Abkürzung für $p \to \wedge$ einführen:

(D_3^0) $\quad p \to \wedge \dashrightarrow \bar{p}$

$(*D_3^0)$ $\quad \bar{p} \dashrightarrow p \to \wedge .$

Als Konsequenzen dieser Definition erhalten wir unmittelbar

(S_{11}^0) $\overline{p \frown \bar{p}}$ $\qquad$ (S_{12}^0) $p_1 \to p_2 \dashrightarrow \bar{p}_2 \to \bar{p}_1$

(S_{13}^0) $p \to \bar{\bar{p}}$ $\qquad$ (S_{14}^0) $\bar{\bar{p}} \leftrightarrow \bar{p}$

(S_{15}^0) $p_1 \frown \bar{p}_2 \to \overline{p_1 \to p_2}$ $\qquad$ (S_{16}^0) $p_1 \frown p_2 \to \overline{p_1 \to \bar{p}_2}$

(S_{17}^0) $\overline{p_1 \frown p_2} \to p_1 \to \bar{p}_2$

(S_{18}^0) $\bar{p}_1 \smile \bar{p}_2 \to \overline{p_1 \frown p_2}$ $\qquad$ (S_{19}^0) $\bigvee\limits_x \overline{p(x)} \to \overline{\bigwedge\limits_x p(x)}$

(S_{20}^0) $\bar{p}_1 \frown \bar{p}_2 \leftrightarrow \overline{p_1 \smile p_2}$ $\qquad$ (S_{21}^0) $\bigwedge\limits_x \overline{p(x)} \leftrightarrow \overline{\bigvee\limits_x p(x)}$.

Die Beweise aller dieser Sätze sind trivial, wenn überall $\bar{p}$ durch $p \to \wedge$ ersetzt wird. Daß hier $\wedge$ für unableitbare Aussagen steht, wird dabei gar nicht benutzt.

Diese Besonderheit von $\wedge$ wird jedoch ausgenützt für

(S_{18}) $\qquad \overline{z_1 = z_2 \frown z_1 \neq z_2}$.

Beweis durch Induktion nach (z_1, z_2):

$\neg | \neq | \to$ 1) $| = | \frown | \neq | \to \wedge$

$\neg z | = | \to$ 2) $z | = | \frown z | \neq | \to \wedge$

$\neg | = z | \to$ 3) $| = z | \frown | \neq z | \to \wedge$

$*D_2, *D_5 \to$ 4) $z_1 = z_2 \frown z_1 \neq z_2 \to \wedge \dashrightarrow z_1 | = z_2 | \frown z_1 | \neq z_2 | \to \wedge$

$S_1 \vdash 1, \ldots, 4 \to S_{18}$.

Wegen

$$t_1 = t_2 \frown t_1 \neq t_2 \rightarrow \bigvee_{z_1} \bigvee_{z_2} (z_1 = z_2 \frown z_1 \neq z_2)$$

folgt weiter

$$(S_{19}) \qquad \frac{}{t_1 = t_2 \frown t_1 \neq t_2}.$$

Die bisher begründete Logik stimmt mit dem sog. Minimalkalkül[4]) überein. Dieser unterscheidet sich vom sog. intuitionistischen Kalkül dadurch, daß

$$(S^0_{22}) \qquad \land \rightarrow p$$

nicht benutzt wird. Wir werden aber jetzt diese Regel durch Induktion über alle Aussagen p beweisen.

(1) $$\land \rightarrow z_1 = z_2.$$

Beweis durch Induktion nach (z_1, z_2):

$$D_1 \rightarrow 1) \land \rightarrow \mathsf{I} = \mathsf{I}$$
$$\neg z\mathsf{I} = \mathsf{I} \rightarrow 2) \land \rightarrow z\mathsf{I} = \mathsf{I}$$
$$\neg \mathsf{I} = z\mathsf{I} \rightarrow 3) \land \rightarrow \mathsf{I} = z\mathsf{I}$$
$$D_2 \rightarrow 4) \land \rightarrow z_1 = z_2 \rightarrow \land \rightarrow z_1 \neq z_2.$$

(2) $$\land \rightarrow z_1 \neq z_2.$$

Beweis durch Induktion nach (z_1, z_2):

$$\neg \mathsf{I} \neq \mathsf{I} \rightarrow 1) \land \rightarrow \mathsf{I} \neq \mathsf{I}$$
$$D_3 \rightarrow 2) \land \rightarrow z\mathsf{I} \neq \mathsf{I}$$
$$D_4 \rightarrow 3) \land \rightarrow \mathsf{I} \neq z\mathsf{I}$$
$$D_5 \rightarrow 4) \land \rightarrow z_1 \neq z_2 \rightarrow \land \rightarrow z_1\mathsf{I} \neq z_2\mathsf{I}.$$

(3) $$\land \rightarrow t_1 = t_2.$$

Beweis.

$$D_{14}, \land \rightarrow z_1 = z_2 \rightarrow 1)\ t_1 = z_1 \frown t_2 = z_2 \frown \land \rightarrow t_1 = t_2$$
$$S_{14} \rightarrow 2) \bigvee_{z_1} \bigvee_{z_2} (t_1 = z_1 \frown t_2 = z_2)$$
$$1, 2 \rightarrow 3) \land \rightarrow t_1 = t_2.$$

(4) $$\land \rightarrow t_1 \neq t_2.$$

Beweis analog zu (3).

[4]) Vgl. hierzu und zum intuitionistischen Kalkül HILBERT-BERNAYS, Grundlagen d. Math. II.

Damit ist (S°_{22}) für jede Primaussage bewiesen. Nach den Regeln für die Operationen $\frown$, $\smile$, $\bigwedge_x$, $\bigvee_x$ folgt dann (S°_{22}) auch für alle mit ihnen zusammengesetzten Aussagen. Da aus $\wedge \to p_2$ auch $\wedge \dashrightarrow p_1 \to p_2$ folgt, gilt (S°_{23}) für jede Aussage.

Wir haben jetzt den intuitionistischen Kalkül erreicht. Daher sind die folgenden Entwicklungen bereits im wesentlichen bekannt. Mit (S°_{23}) ergeben sich als weitere Konsequenzen

(S°_{23}) $\quad p_1 \frown \bar{p}_1 \to p_2$

(S°_{24}) $\quad p_1 \smile p_2 \dashrightarrow \bar{p}_1 \to p_2$

(S°_{25}) $\quad \bar{p}_1 \smile p_2 \dashrightarrow p_1 \to p_2$.

Zwei Aussagen p_1, p_2 heißen komplementär, wenn

(1) $\quad p_1 \smile p_2$ $\qquad$ (2) $\quad \overline{p_1 \frown p_2}$

gelten. Aus (S°_{24}) und (1) folgt $\bar{p}_1 \to p_2$ und $\bar{p}_2 \to p_1$, aus (2) folgt $p_1 \to \bar{p}_2$ und $p_2 \to \bar{p}_1$. Komplementäre Aussagen negieren sich also gegenseitig. Gibt es zu p eine komplementäre Aussage, dann ist diese äquivalent mit $\bar{p}$. Es gibt also genau dann ein Komplement zu p, wenn das tertium non datur

I. $\quad p \smile \bar{p}$

gilt. Wegen (S°_{24}) folgt aus I.

II. $\quad \bar{\bar{p}} \to p$.

Wegen (S_{17}), (S_{19}) gilt I., also auch II. für jede Primaussage. Die Gültigkeit von I., II. ist aber nicht auf Primaussagen beschränkt. Die Gültigkeit von I. läßt sich auf alle „kleinen" Aussagen (d. h. Aussagen ohne die großen Operatoren: Generalisation, Partikularisation) ausdehnen. Der Beweis läßt sich durch Induktion über alle kleinen Aussagen führen. Aus $p_1 \smile \bar{p}_1$ und $p_2 \smile \bar{p}_2$ folgt nämlich zunächst

(1) $\quad (p_1 \frown p_2) \smile \bar{p}_1 \smile \bar{p}_2$

(2) $\quad p_1 \smile p_2 \smile (\bar{p}_1 \frown \bar{p}_2)$

(3) $\quad \bar{p}_1 \smile p_2 \smile (p_1 \frown \bar{p}_2)$

und dann mit (S°_{14}), (S°_{20}), (S°_{22}) und (S°_{15})

(1′) $\quad (p_1 \frown p_2) \smile \overline{p_1 \frown p_2}$

(2′) $\quad (p_1 \smile p_2) \smile \overline{p_1 \smile p_2}$

(3′) $\quad (p_1 \to p_2) \smile \overline{p_1 \to p_2}$.

Die Gültigkeit von II. läßt sich außerdem auf alle „disjunktionsfreien" Aussagen (d. h. Aussagen, die ohne $\smile$, $\bigvee_x$, kleine und große

Disjunktion, zusammengesetzt sind) ausdehnen. Aus $\bar{p}_1 \to p_1; \bar{\bar{p}}_2 \to p_2$ folgt nämlich wegen

$$\overline{\overline{p_1 \frown p_2}} \to \bar{\bar{p}}_1; \quad \overline{\overline{p_1 \frown p_2}} \to \bar{\bar{p}}_2$$

sofort $p_1 \frown p_2 \to p_1 \frown p_2$.
Ebenso folgt aus

$$\overline{\overline{p(x)}} \underset{x}{\to} p(x)$$

sofort $\overline{\overline{\underset{x}{\frown} p(x)}} \to \underset{x}{\frown} p(x)$.

Aus $\bar{p}_2 \to p_2$ folgt schließlich $\overline{\overline{p_1 \to \bar{p}_2}} \to p_1 \to p_2$ (S^{∞}_{11}), (S^{∞}_{12}). Nach diesem Beweis gilt II. noch über den Bereich der disjunktionsfreien Formeln hinaus. Damit II. für $p_1 \to p_2$ gilt, braucht es ja nur für p_2 zu gelten.

Da $\frown$ unableitbar ist, folgt aus der Äquivalenz aller unableitbaren Aussagen

$$\neg p \to \bar{p}.$$

Um auch die Umkehrung

$$\bar{p} \to \neg p$$

zu beweisen, betrachten wir außer den arithmetischen Aussagen p, $p_1, \ldots$ noch die semiotischen Aussagen $p_1 \equiv p_2$, $p_1 \not\equiv p_2$. Diese ergeben sich — wie in der Arithmetik — leicht als komplementär.

(S^{∞}_{11}) $\quad x_1 \equiv x_2 \vee x_1 \not\equiv x_2$

(S^{∞}_{12}) $\quad \overline{x_1 \equiv x_2 \frown x_1 \not\equiv x_2}.$

Wir definieren eine Funktion

$$f(p_1, p_2)$$

durch

$(D^{\infty}_{2,1})$ $\quad p_1 \equiv p_2 \to f(p_1, p_2) \equiv \frown$

$(D^{\infty}_{2,2})$ $\quad p_1 \not\equiv p_2 \to f(p_1, p_2) \equiv p_2.$

Für diese überall definierte Funktion gilt

(S^{∞}_{13}) $\quad \bar{p}_1, p_2 \to f(p_1, p_2).$

Beweis.

$S^{\infty}_1 \to$ 1) $p_1 \equiv p_2, \bar{\bar{p}}_1, p_2 \to \frown$

1, $D^{\infty}_{2,1} \to$ 2) $p_1 \equiv p_2, \bar{\bar{p}}_1, p_2 \to f(p_1, p_2)$

3) $p_1 \not\equiv p_2, \bar{\bar{p}}_1, p_2 \to p_2$

$D^{\infty}_{2,2}$, 3 $\to$ 4) $p_1 \not\equiv p_2, \bar{\bar{p}}_1, p_2 \to f(p_1, p_2)$

S^{∞}_{11}, 2, 4 $\to S^{\infty}_{13}$.

Nach Definition der Variablen $s, s_1, \ldots$ als Variable für die arithmetischen Sätze gilt für jedes Aussagenprädikat $P(p)$

$$p, \underset{s}{P(s)} \to P(p).$$

Wegen der Unableitbarkeit von $\wedge$, d. h. $s \underset{s}{\not\equiv} \wedge$ folgt hieraus

$$(S_{14}^{\infty}) \qquad p \to p \not\equiv \wedge.$$

Jetzt ergibt sich

$$(S_{15}^{\infty}) \qquad \bar{p} \to p \not\equiv s.$$

Beweis.

$$\begin{aligned} & 1)\ s \\ S_{13}^{\infty},\ 1 \to & 2)\ \bar{p} \to f(p, s) \\ S_{14}^{\infty},\ 2 \to & 3)\ \bar{p} \to f(p, s) \not\equiv \wedge \\ D_{9,1}^{\infty} \to & 4)\ p \equiv s \to f(p, s) \equiv \wedge \\ S_{11}^{\infty},\ S_{13}^{\infty},\ 4 \to & 5)\ f(p, s) \not\equiv \wedge \to p \not\equiv s \\ & 3, 5 \to S_{15}^{\infty}. \end{aligned}$$

Damit ist $\neg p \leftrightarrow \bar{p}$ bewiesen. Erst dadurch sind umgangssprachliche Überlegungen wie: „$p_1 \to p_2$ sei bewiesen und p_2 sei unableitbar — dann ist p_1 unableitbar — denn, wäre p_1 ableitbar, so ergäbe sich mit $p_1 \to p_2$ eine Ableitung von p_2" begründet.

Da „unableitbar" mit der Negation von „ableitbar" äquivalent ist, handelt es sich um eine Einkleidung von (S_{12}^{∞}). „Unableitbar" und „ableitbar" sind dagegen nicht komplementär[5]).

§ 9.

Tertium non datur.

Die asymmetrischen Verhältnisse, die in den Sätzen von § 8 zum Ausdruck kommen, legen es nahe, die Untersuchung von Ableitbarkeitsfragen dadurch zu erleichtern, daß man die Schwierigkeit zerlegt: indem zunächst so getan wird, als ob die Verhältnisse symmetrisch wären — und erst nachträglich in den Fällen, in denen sie es nicht sind, die Asymmetrie zusätzlich berücksichtigt wird. Um diese Zerlegung durchzuführen, werden wir jeder Aussage p eine Aussage p' zuordnen, derart, daß (1) für die Ableitbarkeit von p' nur symmetrische Regeln in Betracht zu ziehen sind, und (2) in vielen Fällen $p' \to p$ trivial ist, also nur in einigen Fällen eine zusätzliche Untersuchung erforderlich ist.

[5]) Zum Beweis der Komplementarität wäre ein allgemeines Entscheidungsverfahren erforderlich. Vgl. dazu Hilbert-Bernays, Grundlagen d. Math. II.

Wir definieren p' folgendermaßen:

$$\begin{aligned} p' &\equiv p \quad \text{für Primaussagen} \\ (p_1 \frown p_2)' &\equiv p_1' \frown p_2' \\ (\underset{x}{\frown} p(x))' &\equiv \underset{x}{\frown} p'(x) \\ (p_1 \to p_2)' &\equiv p_1' \to p_2' \\ (p_1 \smile p_2)' &\equiv \overline{\overline{p_1'} \smile \overline{p_2'}} \\ (\underset{x}{\vee} p(x))' &\equiv \overline{\underset{x}{\vee} \overline{p'(x)}}. \end{aligned}$$

Die Abbildung von p auf p' ist also homomorph bezügl. der Operationen $\to$, $\frown$, $\underset{x}{\frown}$, nur bei den Disjunktionen kommt eine doppelte Negation hinzu.

Aus der Definition folgt wie für disjunktionsfreie Aussagen[6])

$$(S^0_{20}) \qquad \overline{\overline{p'}} \to p'.$$

Wir untersuchen jetzt Ableitungen nach den bisherigen Regeln $D^0_1 - D^0_5$, $^*D^0_1 - {}^*D^0_5$, zu denen wir noch das tertium non datur

$$\text{I.} \qquad p \smile \bar{p}$$

hinzunehmen. Da I. nicht allgemein gilt, nennen wir solche Ableitungen „fiktiv" und nennen die bisherigen Ableitungen zum Unterschied „konstruktiv".

1. Satz. *Ist p fiktiv ableitbar, so ist p' konstruktiv ableitbar.*

Zum Beweis ist nur zu zeigen, daß jede Regel der fiktiven Ableitungen durch die Abbildung von p auf p' in eine konstruktivgültige Regel übergeht. Nehmen wir das tertium non datur voraus, das in

$$\overline{\overline{p' \smile \bar{p}'}}$$

(Beweis: $S^0_{20} \vdash S^0_{11} \to \overline{\overline{p \smile \bar{p}}}$) übergeht, so bleiben zunächst die Regeln über die Konjunktionen. Bei diesen ist wegen der Homomorphie der Abbildung nichts zu beweisen. Die Regeln über die Disjunktionen gehen über in

(1) $$p_1' \to \overline{\overline{p_1' \smile p_2'}}.$$

Beweis trivial.

(2) $$p_2' \to \overline{\overline{p_1' \smile p_2'}}.$$

Beweis trivial.

(3) $$p_1' \to p';\ p_2' \to p' \mathrel{\dot{-}} \overline{\overline{p_1' \smile \overline{p'}}} \to p'$$

[6]) Vgl. II in § 8.

Beweis.

$$S^{\circ}_{12} \to \ 1)\ p'_1 \to p' \doteq \overline{p'} \to \overline{p'_1}$$

$$S^{\circ}_{12} \to \ 2)\ p'_2 \to p' \doteq \overline{p'} \to \overline{p'_2}$$

$$S^{\circ}_{3},\ 1,\ 2 \to \ 3)\ p'_1 \to p';\ p'_2 \to p' \doteq \overline{p'} \to \overline{p'_1} \frown \overline{p'_2}$$

$$S^{\circ}_{12},\ 3 \to \ 4)\ p'_1 \to p';\ p'_2 \to p' \doteq \overline{\overline{p'_1} \frown \overline{p'_2}} \to \overline{p'}$$

$$S^{\circ}_{20},\ S^{\circ}_{20},\ 4 \to \ 5)\ p'_1 \to p';\ p'_2 \to p' \doteq \overline{\overline{p'_1} \smile \overline{p'_2}} \to p'.$$

(4) $$p'(x) \to \overline{\overline{\bigvee_x p'(x)}}.$$

Beweis trivial.

(5) $$p'(x) \underset{x}{\to} {}^{(x)}p'_0 \doteq \overline{\overline{\bigvee_x p_i(x)}} \to {}^{(x)}p'_0.$$

Beweis.

$$S^{\circ}_{12} \to \ 1)\ p'(x) \underset{x}{\to} {}^{(x)}p'_0 \doteq \overline{{}^{(x)}p'_0} \to \overline{p'(x)}$$

$$S^{\circ}_{3},\ \underset{x}{1} \to \ 2)\ p'(x) \underset{x}{\to} {}^{(x)}p'_0 \doteq \overline{{}^{(x)}p'_0} \to \bigwedge_x \overline{p'(x)}$$

$$S^{\circ}_{12},\ 2 \to \ 3)\ p'(x) \underset{x}{\to} {}^{(x)}p'_0 \doteq \overline{\bigwedge_x \overline{p'(x)}} \to {}^{(x)}p'_0$$

$$S^{\circ}_{21},\ S^{\circ}_{20},\ 3 \to \ 4)\ p'(x) \underset{x}{\to} {}^{(x)}p'_0 \doteq \overline{\overline{\bigvee_x p'(x)}} \to {}^{(x)}p'_0.$$

Dieses Resultat gestattet jetzt in gewissen Fällen aus der fiktiven Ableitbarkeit von p auch auf die konstruktive Ableitbarkeit zu schließen, nämlich dann, wenn p eine kleine Aussage ist, und dann wenn p disjunktionsfrei ist.

2. Satz. *Ist p eine kleine Aussage oder disjunktionsfrei, so gilt*

$$p' \to p.$$

Beweis. Für disjunktionsfreie Aussagen gilt $p' \equiv p$.

Um den Satz auch für kleine Aussagen zu beweisen, genügt es, für kleine Aussagen p_1, p_2 die Gültigkeit von

$$p'_1 \to p_1;\ p'_2 \to p_2 \doteq (p_1 \smile p_2)' \to p_1 \smile p_2$$

zu zeigen.

$$D^{\circ}_{3},\ {}^{*}D^{\circ}_{3} \to \ 1)\ p'_1 \to p_1;\ p'_2 \to p_2 \doteq p'_1 \smile p'_2 \to p_1 \smile p_2$$

$$\text{II.},\ 1 \to \ 2)\ p'_1 \to p_1;\ p'_2 \to p_2 \doteq \overline{p'_1} \smile \overline{p'_2} \to p_1 \smile p_2.$$

In diesen beiden Sätzen ist insbesondere die Widerspruchsfreiheit der fiktiven Ableitbarkeit enthalten. Wäre $p \frown \bar{p}$ fiktiv ableitbar, so auch $\wedge$. $\wedge$ gehört zu den in Satz 2 genannten Aussagen. Wäre daher $\wedge$ fiktiv ableitbar, so wäre $\wedge$ auch konstruktiv ableitbar. $\wedge$ ist aber nicht konstruktiv ableitbar, folglich ist $p \frown \bar{p}$ auch nicht fiktiv ableitbar.

Für die Aussagen, die nicht zu den in Satz 2 genannten Aussagen gehören, kann aus der fiktiven Ableitbarkeit von p zunächst nichts anderes als die konstruktive Ableitbarkeit von p' gefolgert werden. Wir fügen jedoch hinzu:

3. Satz. Für implikations- und generalisationsfreie Aussagen gilt

$$p' \to p .$$

Beweis durch Induktion über alle Aussagen.

Für Primaussagen ergibt sich Satz 3 aus $p' \equiv p$ (S°_{13}).

Aus $p_1' \to \bar{\bar{p}}_1$; $p_2' \to \bar{\bar{p}}_2$ folgt

(1) $$(p_1 \frown p_2)' \to \overline{\overline{p_1 \frown p_2}} .$$

Beweis.

$$\begin{aligned} p_1' \to \bar{\bar{p}}_1,\ p_2' \to \bar{\bar{p}}_2 \to 1)\ & p_1' \frown p_2' \to \bar{\bar{p}}_1 \frown \bar{\bar{p}}_2 \\ S^{\circ}_{15},\ S^{\circ}_{17} \to 2)\ & \bar{\bar{p}}_1 \frown \bar{\bar{p}}_2 \to \overline{\overline{p_1 \frown p_2}} \\ 1, 2 \to 3)\ & p_1' \frown p_2' \to \overline{\overline{p_1 \frown p_2}} . \end{aligned}$$

(2) $$(p_1 \smile p_2)' \to \overline{\overline{p_1 \smile p_2}} .$$

Beweis.

$$\begin{aligned} p_1' \to \bar{\bar{p}}_1;\ p_2' \to \bar{\bar{p}}_2 \to 1)\ & p_1' \smile p_2' \to \bar{\bar{p}}_1 \smile \bar{\bar{p}}_2 \\ S^{\circ}_{12} \to 2)\ & \bar{\bar{p}}_1 \smile \bar{\bar{p}}_2 \to \overline{\overline{p_1 \smile p_2}} \\ 1, 2 \to 3)\ & p_1' \smile p_2' \to \overline{\overline{p_1 \smile p_2}} \\ S^{\circ}_{12},\ S^{\circ}_{14} \to 4)\ & \overline{\overline{p_1' \smile p_2'}} \to \overline{\overline{p_1 \smile p_2}} . \end{aligned}$$

Aus $p'(x) \underset{x}{\to} \overline{\overline{p(x)}}$ folgt

(3) $$(\underset{x}{\vee}\, p(x))' \to \overline{\overline{\underset{x}{\vee}\, p(x)}} .$$

Beweis.

$$\begin{aligned} p'(x) \underset{x}{\to} \overline{\overline{p(x)}} \to 1)\ & \underset{x}{\vee}\, p'(x) \to \underset{x}{\vee}\, \overline{\overline{p(x)}} \\ S^{\circ}_{12} \to 2)\ & \underset{x}{\vee}\, \overline{\overline{p(x)}} \to \underset{x}{\vee}\, p(x) \\ 1, 2 \to 3)\ & \underset{x}{\vee}\, p'(x) \to \underset{x}{\vee}\, p(x) \\ S^{\circ}_{12}, S^{\circ}_{14}, 3 \to 4)\ & \overline{\overline{\underset{x}{\vee}\, p'(x)}} \to \overline{\overline{\underset{x}{\vee}\, p(x)}} . \end{aligned}$$

Die fiktiven Ableitungen sind die Ableitungen mithilfe der sog. klassischen, zweiwertigen Logik, denn für diese sind die folgenden Negationsregeln zulässig:

$$p \smile \bar{p},\quad \overline{p \frown \bar{p}},\quad \bar{\bar{p}} \leftrightarrow p,$$

$$\overline{p_1 \frown p_2} \leftrightarrow \bar{p}_1 \smile \bar{p}_2,\quad \overline{p_1 \smile p_2} \leftrightarrow \bar{p}_1 \frown \bar{p}_2,$$

$$\overline{\underset{x}{\wedge}\, p(x)} \leftrightarrow \underset{x}{\vee}\, \bar{p}(x),\quad \underset{x}{\vee}\, p(x) \leftrightarrow \underset{x}{\wedge}\, \bar{p}(x),$$

$$p_1 \to p_2 \leftrightarrow \bar{p}_1 \smile p_2 .$$

Die letzte Äquivalenz, die gestattet, jede Regel implikationsfrei zu schreiben, (wobei die Negation nach den vorhergehenden Äquivalenzen eliminiert wird) macht es verständlich, daß man glaubte, die Konseqenzlogik durch die zweiwertige Logik entbehrlich machen zu können.

Eine Aussage, wie $p_1 \to p_2 \dashrightarrow p_1 \dashrightarrow p_1$ ist aber zwar unter Voraussetzung von $p_1 \sim \bar{p}_1$ beweisbar, also fiktiv ableitbar, sie ist aber kein konsequenzlogischer Satz.

Der hier eingeschlagene Weg zeigt jedoch, daß die zweiwertige Logik eine einfache (trotz der bisherigen Kompliziertheit der Widerspruchsfreiheitsbeweise) und sinnvolle Erweiterung der positiven Logik ist.

(Eingegangen am 13. Februar 1950.)